


```

; PRIOR APPLICATION NUMBER: US5N 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US5N 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: US5N 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US5N 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-679-25

```

Query Match	18.3%	Score 395;	DB 9;	Length 366;
Best Local Similarity	30.5%	Pred. No. 1.4e-24;		
Matches 106;	Conservative	48;	Mismatches 132;	Indels 62;
Gaps				13;

[illegible]

```

RESULT 28
US-09-943-718-4
; Sequence 4, Application US/09943718
; Patent No. US20020103361A1
; GENERAL INFORMATION:
; APPLICANT: Huffine, Constance F.
; Rossi, Devora L.
; Capone, Myrtam
; Hedrick, Joseph A.
; Vicari, Alain
; Gorman, Daniel M.
; Zlotnik, Albert
; TITLE OF INVENTION: Mammalian Chemokines; Receptors;
;
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

, APPLICATION NUMBER: US/09/943,718
, FILING DATE: 30-AUG-2001
, CLASSIFICATION: <unknown>
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 09/009,817
, FILING DATE: 20-APR-1998
, ATTORNEY/AGENT INFORMATION:
, NAME: Ching, Edwin P.
, REGISTRATION NUMBER: 34,090
, REFERENCE/DOCKET NUMBER: DX0589K
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (650)852-9196
, TELEFAX: (650)496-1200
, INFORMATION FOR SEQ ID NO: 4:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 360 amino acids
, TYPE: amino acid
, TOPOLOGY: linear
, MOLECULE TYPE: protein
, SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-943-718-4

```

Query Match	17.8%;	Score	384.5;	DB	10;	Length	360;
Best Local Similarity	35.7%;	Pred. No.	9.6e-24;				
Matches	85;	Conservative	38;	Mismatches	96;	Indels	19;
Gaps	4;						
Qy	153	NLAIDPFCVTPFKIAVHLNGNNVFGVLCRATTIVFYGNMYCSILLACISINRYL	212				
Db	93	NLAVDLLALVLPRLAYHLRGQRPFGAAACRVATAALYGHMYSVLLLAASLDRYL	152				
Qy	213	AIVHPFTYRGLPKHTYALVTTCGLVWATVFLYMLPFFLIKOEYVLVQPDITTCCHDVHNTC-	271				
Db	153	ALVHPRARALRGORTTGTGCLVAWLSAAATLPLTLHRQNFRLAP-IACC--VMRCP	209				
Qy	272	-----ESSFPQLYFYFISLAFFGLIPFVLLIYCYAAIIRTLNAYDRHLWYVKASL	323				
Db	210	WLSRTPGTGPPSAWLSWAASL-----PLLAWGLCYGTTLRALAANGORYSHALRLTA	262				
Qy	324	LILVIETICFAPSNIILIIHHANYYYNNNTDGLFYIYLIALCLGSLNSCLDPFLYFLMS	381				
Db	263	LVLFSVASFTPSNVLLVLYHSNPSPEAMGNLYGAVYPSLALSTLNSCVDPFIYVYVS	320				

RESULT 29
 US-09-077-173A-2
 ; Sequence 2, Application US/09077173A
 ; Publication No. US20030082674A1
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: RECEPTOR AND NUCLEIC ACID MOLECULE ENCODING
 ; TITLE OF INVENTION: SAID RECEPTOR
 ; NUMBER OF SEQUENCES: 4
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/077,173A
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/BE 96/00123
 ; FILING DATE:
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 365 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-077-173A-2

Query Match 17.7%; Score 382; DB 9; Length 365;
Best Local Similarity 31.8%; Pred. No. 1.6e-23;

Db 183 SVTSCLELN--LYKIAKLQTMNYIALV-VGCLLPFTLSICYLLIIRVLLKVEVPESGLR 239
QY 313 --HRWLWYKASLLIIVFTICFAPSNIILIIHHANYYN-NTDGLYFYIYALICLGSLN 369
Db 240 VSHRKA--LTTIIITLIIFCLFPLPYHTLRTVHLTWKVLCKDLRHLKALVITLALAAAN 297
QY 370 SCLDPELYF 378
Db 298 ACFNPLLY 306
RESULT 32
US-09-779-679-26
; Sequence 26, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sam
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779, 679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-679-26
Query Match 17.5%; Score 376.5; DB 9; Length 346;
Best Local Similarity 30.1%; Pred. No. 4.1e-23;
Matches 93; Conservative 62; Mismatches 129; Indels 25; Gaps 9;
QY 82 TITVKIKCPESASHLHVKNATMGYLTSSLTKLIPAIYLLVFGVGPANAVTLMLFFR 141
Db 11 SISVSEMEPNGTFSNNRNCII---ENFKREFFPIVLIIFFWGLGSLSIY-VFLQ 65
QY 142 --TRSICTTFTYNTLAIDFLFCVTLPPKIAIYHLNGNNWVGEVLCPATTVIPYGNMYS 199
Db 66 PYKSTSVNVFVLMNLAIISDLTSLPFRADYIRGNSWIFGDLACRIMSISLYVNMYS 125
QY 200 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLWATVFLYMLPFFILKQEYVLVQ 259
Db 126 IYFLTVLSVVRFLAMVHPRLHVTIRSAILCGIIM---ILIMASSIMLLDSGSEQNG 182
QY 260 DITTDHVNTESSSPFQLYYFISLAFEGFLIPFVLIYCYAAIIRTNAYD-----312
Db 183 SVTSCLELN--LYKIAKLQTMNYIALV-VGCLLPFTLSICYLLIIRVLLKVEVPESGLR 239
QY 313 --HRWLWYKASLLIIVFTICFAPSNIILIIHHANYYN-NTDGLYFYIYALICLGSLN 369
Db 240 VSHRKA--LTTIIITLIIFCLFPLPYHTLRTVHLTWKVLCKDLRHLKALVITLALAAAN 297
QY 370 SCLDPELYF 378
Db 298 ACFNPLLY 306
RESULT 34
US-09-866-230-7
; Sequence 7, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.

Db 126 IYFLTVLSVVRFLAMVHPRLHVTIRSAILCGIIM---ILIMASSIMLLDSGSEQNG 182
QY 260 DITTDHVNTESSSPFQLYYFISLAFEGFLIPFVLIYCYAAIIRTNAYD-----312
Db 183 SVTSCLELN--LYKIAKLQTMNYIALV-VGCLLPFTLSICYLLIIRVLLKVEVPESGLR 239
QY 313 --HRWLWYKASLLIIVFTICFAPSNIILIIHHANYYN-NTDGLYFYIYALICLGSLN 369
Db 240 VSHRKA--LTTIIITLIIFCLFPLPYHTLRTVHLTWKVLCKDLRHLKALVITLALAAAN 297
QY 370 SCLDPELYF 378
Db 298 ACFNPLLY 306
RESULT 33
US-09-826-791-6
; Sequence 6, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1el Polypeptide
; FILE REFERENCE: PC10914ADAM
; CURRENT APPLICATION NUMBER: US/09/826,791
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 0008504.3
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/198,367
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-791-6
Query Match 17.5%; Score 376.5; DB 10; Length 346;
Best Local Similarity 30.1%; Pred. No. 4.1e-23;
Matches 93; Conservative 62; Mismatches 129; Indels 25; Gaps 9;
QY 82 TITVKIKCPESASHLHVKNATMGYLTSSLTKLIPAIYLLVFGVGPANAVTLMLFFR 141
Db 11 SISVSEMEPNGTFSNNRNCII---ENFKREFFPIVLIIFFWGLGSLSIY-VFLQ 65
QY 142 --TRSICTTFTYNTLAIDFLFCVTLPPKIAIYHLNGNNWVGEVLCPATTVIPYGNMYS 199
Db 66 PYKSTSVNVFVLMNLAIISDLTSLPFRADYIRGNSWIFGDLACRIMSISLYVNMYS 125
QY 200 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLWATVFLYMLPFFILKQEYVLVQ 259
Db 126 IYFLTVLSVVRFLAMVHPRLHVTIRSAILCGIIM---ILIMASSIMLLDSGSEQNG 182
QY 260 DITTDHVNTESSSPFQLYYFISLAFEGFLIPFVLIYCYAAIIRTNAYD-----312
Db 183 SVTSCLELN--LYKIAKLQTMNYIALV-VGCLLPFTLSICYLLIIRVLLKVEVPESGLR 239
QY 313 --HRWLWYKASLLIIVFTICFAPSNIILIIHHANYYN-NTDGLYFYIYALICLGSLN 369
Db 240 VSHRKA--LTTIIITLIIFCLFPLPYHTLRTVHLTWKVLCKDLRHLKALVITLALAAAN 297
QY 370 SCLDPELYF 378
Db 298 ACFNPLLY 306
RESULT 34
US-09-866-230-7
; Sequence 7, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.

; TITLE OF INVENTION: No. US20020150901a1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230-7

Query Match 17.5%; Score 376.5; DB 10; Length 346;
Best Local Similarity 30.1%; Pred. No. 4.1e-23;
Matches 93; Conservative 62; Mismatches 129; Indels 25; Gaps 9;

QY 82 TITVKIKCPESASHLHVKNATMGYLTSSLTSLKLIPIAYLLVFGVGPANAVTLMMLFFR 141
DB 11 SISVSEMEPGTFTSNNSRNTCTI----ENFKREFPIVYLIIFPWGLGNLSIY-VFLQ 65

QY 142 --TRSICTTVFYTNLAIDELFCVTLPFKIAYHLGNWVFGVLCRATTIVIFYGNMYS 199
DB 66 PYKSTSVNVFVMLNLAISDLFLSTLPFRADYLRGSNWIFGDLACRIMSYSLVNMYS 125

QY 200 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLWATVFLYMLPFFILKOEYLVQP 259
DB 126 IYFLTVLSVVRFLAMVHPFRLHVTIRSASWILGGIIV---ILIMASSIMLLDSGSEQNG 182

QY 260 DITTDVHTNCTSSSPFQYFISLAFGLFPLFVLIYCYAAIIRTLNAYD----- 312
DB 183 SVTSCLELN--LYKIAKLQTMNYIALV-VGCLLPFFTLSCYLLIIRVLLKVEVPESGLR 239

QY 313 --HRWLWVVKASLLILVIFTICFAPSNIILIIHHANYYN-NTDGLYFIYLIACLGSLN 369
DB 240 VSHRKA--LTTIITLIIFLCFLPYHTLRTVHLTTWKVGLCKDRLHKALVITLALAAAN 297

QY 370 SCLDPFLYF 378
DB 298 ACFNPLLYY 306

RESULT 35
US-09-866-230-9
; Sequence 9, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901a1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230-9

Query Match 17.5%; Score 376.5; DB 10; Length 346;
Best Local Similarity 30.1%; Pred. No. 4.1e-23;
Matches 93; Conservative 62; Mismatches 129; Indels 25; Gaps 9;

QY 82 TITVKIKCPESASHLHVKNATMGYLTSSLTSLKLIPIAYLLVFGVGPANAVTLMMLFFR 141
DB 11 SISVSEMEPGTFTSNNSRNTCTI----ENFKREFPIVYLIIFPWGLGNLSIY-VFLQ 65

QY 142 --TRSICTTVFYTNLAIDELFCVTLPFKIAYHLGNWVFGVLCRATTIVIFYGNMYS 199

DB 66 PYKSTSVNVFVMLNLAISDLFLSTLPFRADYLRGSNWIFGDLACRIMSYSLVNMYS 125
QY 200 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLWATVFLYMLPFFILKOEYLVQP 259
DB 126 IYFLTVLSVVRFLAMVHPFRLHVTIRSASWILGGIIV---ILIMASSIMLLDSGSEQNG 182
QY 260 DITTDVHTNCTSSSPFQYFISLAFGLFPLFVLIYCYAAIIRTLNAYD----- 312
DB 183 SVTSCLELN--LYKIAKLQTMNYIALV-VGCLLPFFTLSCYLLIIRVLLKVEVPESGLR 239
QY 313 --HRWLWVVKASLLILVIFTICFAPSNIILIIHHANYYN-NTDGLYFIYLIACLGSLN 369
DB 240 VSHRKA--LTTIITLIIFLCFLPYHTLRTVHLTTWKVGLCKDRLHKALVITLALAAAN 297
QY 370 SCLDPFLYF 378
DB 298 ACFNPLLYY 306

RESULT 36
US-10-024-494-8
; Sequence 8, Application US/10024494
; Publication No. US20030044898A1
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; CAO, LIANG
; ROSEN, CRAIG A.
; TITLE OF INVENTION: Human GPR4 G-Protein Coupled Receptor and
; Nucleotides Encoding Same (As Amended)
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSER: STERNE, KESSLER, GOLDSTEIN & FOX P. L. L. C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/024,494
; FILING DATE: 21-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/465,973
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: PCT/US95/04079
; FILING DATE: 30-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1140001/EKS/HCC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-024-494-8

Query Match 17.4%; Score 375; DB 9; Length 344;
Best Local Similarity 30.9%; Pred. No. 5.5e-23;
Matches 96; Conservative 55; Mismatches 122; Indels 38; Gaps 10;

QY 101 NATMGVLTSSLTSLKLIPIAYLLVFGVGPANAVTLMMLFFRTRISIC-----TTVFYTN 153

Query Match 17.3%; Score 372.5; DB 10; Length 337;
Best Local Similarity 30.2%; Pred. No. 8.5e-23;
Matches 103; Conservative 59; Mismatches 140; Indels 39; Gaps 12;

QY 92 ESASHLHVKNATMGYLTSSTKLIPAIYLLVVGVPANAVTLWMLF--FRTRSICTTV 149
DB 2 DETGNLTVSATCHDIDPRNOVYSTLYSMISVVGFGFVLYVLIKYTHKKS-AFQV 60
QY 150 FYTNLAIADEFVCTLPFKIAYHLNGNNWVFGVLCRATTVIYFGNMYCSILLACISIN 209
DB 61 YMINLAVADLLCVCTLPRLRVVYVHKGIWFGDFCLRLSTYALYVNLVCSIFEMTAMSF 120
QY 210 RYLAIVHPFTYRGLPKHTYALVTCGLWATVFLYMLPFFILKOEYVLVQPDITTDHVDHN 269
DB 121 RCIAIVFPVQINILVTKKARFVCGIWFIVILTSSPFLMAKPQ--KDEKNNTKCFEPQ 178
QY 270 TCSSSPFQIYFISLAFFGLIPFVLIYCYAAIIRTL-----NAYDHRWLWVVKAS 322
DB 179 DNOTKRVHLVHVSU--FVGFIIPFVIIVCYTIILTLKSMKNLSSHK-----KAI 232
QY 323 LLILVI---FTICFAPSNIILIIHHANYNNNT---DGLYFI---YLIALCLGSLNSCLD 373
DB 233 GMIMVVTAAFLVSFMPHIQRTI-HLHFLHNETKPCDSYLRMOKSVVITLSLAASNCPP 291
QY 374 PFLYFL-----MSKTRNH---STAYLTXXXNDLREQQ 403
DB 292 PLYFFSGGNFRKRLSTFRKHSLSSTVTVPRKKASLPEKGE 332

RESULT 40

US-09-826-508-10
; Sequence 10, Application US/09826508
; Patent No. US20010025099A1
; GENERAL INFORMATION:
; APPLICANT: Nabil Elshourbagy
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
; TITLE OF INVENTION: and Polynucleotides
; FILE REFERENCE: GP-70744USB
; CURRENT APPLICATION NUMBER: US/09/826,508
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 358
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-826-508-10

Query Match 17.1%; Score 368.5; DB 10; Length 358;
Best Local Similarity 29.3%; Pred. No. 1.9e-22;
Matches 89; Conservative 64; Mismatches 116; Indels 35; Gaps 9;

QY 100 KNATMGYLTSSTKLIPAIYLLVVGVPANAVTLWMLPFRTRSICTTVFY-TNLAIAD 158
DB 32 KNTT---LHNEFDTIPLVPLIIIFVASILLGLAVN-IFPHIRNKTSTFIYLNKIWAD 87
QY 159 FLFCVTLPEKIAHLNGNNWVFGVLCRATTVIYFGNMYCSILLACISINRYLAIVHPF 218
DB 88 LIMTLTFPPFRIVHDAGPGWPKFCLCRYTSVLFYANMYTSIVFLGLISIDRYLKVKPKP 147
QY 219 TYRGLPKHTYALVTCGLWATVFLYMLPFFILKOEYVLVQPDITTDHVDHNTCESSPPQ 278
DB 148 GDSRMYSITTKVLSVCVWIMAVLSLPNLTNG---QP---TEDNHDCSKLKSPLG 200
QY 279 LYFISLAFFG---FLIPFVLIYCYAAIIRTNAYDHRWL-----WYVKASLLILVI 328
DB 201 VKWHTAVTVNSCLFVAVLVILIGCVIAISRYIHKSSRQFISQSRKRKHQNRVWVAV 260
QY 329 FTICFAPSNIILIIHHANYNNNTDG-----LYFYILIALCLGSLNSCLDPFLYFL 379
DB 261 FFTCFULPHLCRI-----PFTFSLHDLRLDDESAQKILYYCKEITLFLSACNVCLDPIIYFF 316

QY 380 MSKT 383
DB 317 MCRS 320
Search completed: June 24, 2003, 12:18:37
Job time : 34.5399 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:01:21 ; Search time 18.0221 Seconds
(without alignments)
666.102 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIM.....AYLTXXNDLREGQPSQRT 408

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2151	99.7	408	2	US-08-742-440A-6
2	1433.5	66.9	407	2	US-08-742-440A-3
3	606.5	28.1	399	1	US-08-476-000-61
4	606.5	28.1	399	1	US-08-472-840-61
5	606.5	28.1	399	2	US-08-476-976-61
6	606.5	28.1	399	3	US-08-474-410-61
7	606.5	28.1	399	4	US-08-486-673B-61
8	605.5	28.1	395	1	US-08-097-938-2
9	605.5	28.1	395	1	US-08-476-000-2
10	605.5	28.1	395	1	US-08-472-840-2
11	605.5	28.1	395	2	US-08-476-976-2
12	605.5	28.1	395	3	US-08-474-410-2
13	603.5	28.0	395	1	US-08-097-938-5
14	603.5	28.0	395	1	US-08-476-000-5
15	603.5	28.0	395	1	US-08-472-840-5
16	603.5	28.0	395	2	US-08-476-976-5
17	603.5	28.0	395	3	US-08-474-410-5
18	603.5	28.0	395	4	US-08-486-673B-2
19	603.5	28.0	395	4	US-08-486-673B-5
20	585.5	27.1	398	1	US-08-097-938-4
21	585.5	27.1	398	1	US-08-476-000-4
22	585.5	27.1	398	1	US-08-472-840-4
23	585.5	27.1	398	2	US-08-476-976-4
24	585.5	27.1	398	3	US-08-474-410-4
25	585.5	27.1	398	4	US-08-486-673B-4
26	584.5	27.0	398	4	US-08-486-673B-6
27	583	27.0	394	2	US-08-742-440A-8

28	581.5	27.0	398	1	US-08-097-938-6	Sequence 6, Appli
29	581.5	27.0	398	1	US-08-476-000-6	Sequence 6, Appli
30	581.5	27.0	398	1	US-08-472-840-6	Sequence 6, Appli
31	581.5	27.0	398	2	US-08-476-976-6	Sequence 6, Appli
32	581.5	27.0	398	3	US-08-474-410-6	Sequence 6, Appli
33	577.5	26.8	397	4	US-08-486-673B-63	Sequence 63, Appli
34	569.5	26.4	397	1	US-08-476-000-63	Sequence 63, Appli
35	569.5	26.4	397	1	US-08-472-840-63	Sequence 63, Appli
36	569.5	26.4	397	2	US-08-476-976-63	Sequence 63, Appli
37	569.5	26.4	397	3	US-08-474-410-63	Sequence 63, Appli
38	526	24.4	425	1	US-08-097-938-7	Sequence 7, Appli
39	526	24.4	425	1	US-08-476-000-7	Sequence 7, Appli
40	526	24.4	425	1	US-08-472-840-7	Sequence 7, Appli
41	526	24.4	425	2	US-08-476-976-7	Sequence 7, Appli
42	526	24.4	425	3	US-08-474-410-7	Sequence 7, Appli
43	526	24.4	425	4	US-08-486-673B-7	Sequence 7, Appli
44	523	24.2	425	1	US-07-657-769B-69	Sequence 69, Appli
45	523	24.2	425	1	US-07-789-184-220	Sequence 220, App

ALIGNMENTS

RESULT 1
US-08-742-440A-6
; Sequence 6, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
US-08-742-440A-6

Query Match 99.7%; Score 2151; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 2.1e-167;
Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
;
; US-08-742-440A-3

Query Match 66.9%; Score 1443.5; DB 2; Length 407;
Best Local Similarity 71.5%; Pred. No. 6.7e-110;
Matches 264; Conservative 45; Mismatches 59; Indels 1; Gaps 1;

QY 14 GTQVIMKALIPFAAGLLLLPTFCQSGMENDTNLAKPTLPIKTRGAPPNPFEPF 73
DB 11 GSQDIKMKILILVAAGLLFLPVTVCQSGI-NVSDNSAKPTLTIKSFNGGQPTFEFP 69
QY 74 ALEGWGTATITVKIKCPESASHLVKNATMGYLTSSLSSTKLI PAIYLAVFVGV PANAV 133
DB 70 DIEGWGTATITKACPEDSISTLHVNNATIGYLRSSLSSTQVPAIYILFVGVPSNV 129
QY 134 TLMLFFRTRISCTTVFYTNLAIAADFLFCVTLFPFKIAHYHLNGNNWVGEVLCRA 193
DB 130 TLWKLRLTKSISLVIFHTNLAIAADLLFCVTLFPFKIAHYHLNGNNWVGEVLCRI 189
QY 194 GMYCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFF 253
DB 190 GMYCAILLITCMGINRYLATAPFTYQKLPKRSFSLMCGIYVWVFLYMLPFF 249
QY 254 YVLVQPDITTCVDVHNTCESSPPFQLYYFISLAFFGLPIPVLIIVCYAAIIR 313
DB 250 YHLVSEITTCVDVDAESPSPFRFYFVSIAFFGLPIPVLIIVCYAAIIR 309
QY 314 RMLWYKASLLILVITFTICFAPSNIILIIHHANYNNYNTDGLYFYLIACLG 373
DB 310 IWLGYIKAVLLILVITFTICFAPTNIILVIHANYNNYNTDGLYFYLIACLG 369
QY 374 PELYFLMSK 382
DB 370 PELYFVMSK 378

RESULT 3
US-08-476-000-61
; Sequence 61, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-1500
;

1 CSMILOISXRLDGTQVIMKALIPFAAGLLLLPTFCQSGMENDTNLAKPTLPIKTR 60
DB 1 CSMILOISXRLDGTQVIMKALIPFAAGLLLLPTFCQSGMENDTNLAKPTLPIKTR 60
QY 61 GAPNPFEPFSALEGWGTATITVKIKCPESASHLVKNATMGYLTSSLSSTKLI PAI 120
DB 61 GAPNPFEPFSALEGWGTATITVKIKCPESASHLVKNATMGYLTSSLSSTKLI PAI 120
QY 121 LVLVGVGPANAVTLWMLFFRTRISCTTVFYTNLAIAADFLFCVTLFPFKIAHYHL 180
DB 121 LVLVGVGPANAVTLWMLFFRTRISCTTVFYTNLAIAADFLFCVTLFPFKIAHYHL 180
QY 181 GEVLCRAITVIFGNYMVCISILLACISINRYLAIVHPFTYRGLPKHTYALVTCGL 240
DB 181 GEVLCRAITVIFGNYMVCISILLACISINRYLAIVHPFTYRGLPKHTYALVTCGL 240
QY 241 FLYMLPFFILKQBYLVLPDITTCVDVHNTCESSPPFQLYYFISLAFFGLPIPVLI 300
DB 241 FLYMLPFFILKQBYLVLPDITTCVDVHNTCESSPPFQLYYFISLAFFGLPIPVLI 300
QY 301 YAAIIRTNLAYDHRWLVYKASLLILVITFTICFAPSNIILIIHHANYNNYNTDGL 360
DB 301 YAAIIRTNLAYDHRWLVYKASLLILVITFTICFAPSNIILIIHHANYNNYNTDGL 360
QY 361 IALCLGSLNSCLDPFLYFLMSKTRNHSYALTITKXNDLREOGQPSORT 408
DB 361 IALCLGSLNSCLDPFLYFLMSKTRNHSYALTITKXNDLREOGQPSORT 408

RESULT 2
US-08-742-440A-3
; Sequence 3, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 407 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
;
; US-08-742-440A-3

Query Match 66.9%; Score 1443.5; DB 2; Length 407;
Best Local Similarity 71.5%; Pred. No. 6.7e-110;
Matches 264; Conservative 45; Mismatches 59; Indels 1; Gaps 1;

QY 14 GTQVIMKALIPFAAGLLLLPTFCQSGMENDTNLAKPTLPIKTRGAPPNPFEPF 73
DB 11 GSQDIKMKILILVAAGLLFLPVTVCQSGI-NVSDNSAKPTLTIKSFNGGQPTFEFP 69
QY 74 ALEGWGTATITVKIKCPESASHLVKNATMGYLTSSLSSTKLI PAIYLAVFVGV PANAV 133
DB 70 DIEGWGTATITKACPEDSISTLHVNNATIGYLRSSLSSTQVPAIYILFVGVPSNV 129
QY 134 TLMLFFRTRISCTTVFYTNLAIAADFLFCVTLFPFKIAHYHLNGNNWVGEVLCRA 193
DB 130 TLWKLRLTKSISLVIFHTNLAIAADLLFCVTLFPFKIAHYHLNGNNWVGEVLCRI 189
QY 194 GMYCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFF 253
DB 190 GMYCAILLITCMGINRYLATAPFTYQKLPKRSFSLMCGIYVWVFLYMLPFF 249
QY 254 YVLVQPDITTCVDVHNTCESSPPFQLYYFISLAFFGLPIPVLIIVCYAAIIR 313
DB 250 YHLVSEITTCVDVDAESPSPFRFYFVSIAFFGLPIPVLIIVCYAAIIR 309
QY 314 RMLWYKASLLILVITFTICFAPSNIILIIHHANYNNYNTDGLYFYLIACLG 373
DB 310 IWLGYIKAVLLILVITFTICFAPTNIILVIHANYNNYNTDGLYFYLIACLG 369
QY 374 PELYFLMSK 382
DB 370 PELYFVMSK 378

RESULT 3
US-08-476-000-61
; Sequence 61, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-1500
;

1 CSMILOISXRLDGTQVIMKALIPFAAGLLLLPTFCQSGMENDTNLAKPTLPIKTR 60
DB 1 CSMILOISXRLDGTQVIMKALIPFAAGLLLLPTFCQSGMENDTNLAKPTLPIKTR 60
QY 61 GAPNPFEPFSALEGWGTATITVKIKCPESASHLVKNATMGYLTSSLSSTKLI PAI 120
DB 61 GAPNPFEPFSALEGWGTATITVKIKCPESASHLVKNATMGYLTSSLSSTKLI PAI 120
QY 121 LVLVGVGPANAVTLWMLFFRTRISCTTVFYTNLAIAADFLFCVTLFPFKIAHYHL 180
DB 121 LVLVGVGPANAVTLWMLFFRTRISCTTVFYTNLAIAADFLFCVTLFPFKIAHYHL 180
QY 181 GEVLCRAITVIFGNYMVCISILLACISINRYLAIVHPFTYRGLPKHTYALVTCGL 240
DB 181 GEVLCRAITVIFGNYMVCISILLACISINRYLAIVHPFTYRGLPKHTYALVTCGL 240
QY 241 FLYMLPFFILKQBYLVLPDITTCVDVHNTCESSPPFQLYYFISLAFFGLPIPVLI 300
DB 241 FLYMLPFFILKQBYLVLPDITTCVDVHNTCESSPPFQLYYFISLAFFGLPIPVLI 300
QY 301 YAAIIRTNLAYDHRWLVYKASLLILVITFTICFAPSNIILIIHHANYNNYNTDGL 360
DB 301 YAAIIRTNLAYDHRWLVYKASLLILVITFTICFAPSNIILIIHHANYNNYNTDGL 360
QY 361 IALCLGSLNSCLDPFLYFLMSKTRNHSYALTITKXNDLREOGQPSORT 408
DB 361 IALCLGSLNSCLDPFLYFLMSKTRNHSYALTITKXNDLREOGQPSORT 408

RESULT 2
US-08-742-440A-3
; Sequence 3, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 407 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```



```
;
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-000-61

Query Match      28.1%; Score 606.5; DB 1; Length 399;
Best Local Similarity 35.4%; Pred. No. 7e-42;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;

QY 22 ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPNSFEFFPFALEGWT 79
Db 5 SLAWLLGGITLLAASVSCSRNTENLAPGRNNSKGRSLIGRLETQPP-----IT 51
QY 80 GATITVK--IKCPESASHLHVKNATMGVLTSSLSKLIPIAYLLVVFVGVVANATLWM 137
Db 52 GKGVPVEPGFSIDFSAS-----ILTGKLTTFVLPVVYIIIVFVIGLPSNGMALWI 101
QY 138 LPEFTSICITTFY-TNLAIADFLFCVTLPPFKIAYHLGNMNVFGEVLCRAITVIFGNN 196
Db 102 FLFRTKKHPAVIYMANLADLLSVIWFPLKISYHLGNMNVYGEALCKVLIGFYGNN 161
QY 197 YCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKOE 253
Db 162 YCSILFMTCLSVORYVIVNPM---GHPRKK-ANIAVGVSALWLLIFLVTIPLYVMKOT 217
QY 254 YVLVQPDITTCDDVHNTCESSPPQLY-----YFISLAPFGFLIPFVLIICYAAIIRT 307
Db 218 IYIPALNITTCDDV-----LPEEVLVGMDFNYFLSLAIGVLFPAALLTASAYVLMIKT 270
QY 308 LNA-----YDHRWLWYVKASLLIIVFTICFAPSNIILIIHHANYNNNTDGLYFIYLI 361
Db 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNLLLVVHFLIKTORQSHVYALYLV 330
QY 362 ALCLGSLNSCLDPFLYFLMSKT-RNHS 387
Db 331 ALCLSTLNSCIDPFVYFVSKDFRDHA 357

RESULT 4
US-08-476-840-61
; Sequence 61, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
```

```
;
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-840-61

Query Match      28.1%; Score 606.5; DB 1; Length 399;
Best Local Similarity 35.4%; Pred. No. 7e-42;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;

QY 22 ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPNSFEFFPFALEGWT 79
Db 5 SLAWLLGGITLLAASVSCSRNTENLAPGRNNSKGRSLIGRLETQPP-----IT 51
QY 80 GATITVK--IKCPESASHLHVKNATMGVLTSSLSKLIPIAYLLVVFVGVVANATLWM 137
Db 52 GKGVPVEPGFSIDFSAS-----ILTGKLTTFVLPVVYIIIVFVIGLPSNGMALWI 101
QY 138 LPEFTSICITTFY-TNLAIADFLFCVTLPPFKIAYHLGNMNVFGEVLCRAITVIFGNN 196
Db 102 FLFRTKKHPAVIYMANLADLLSVIWFPLKISYHLGNMNVYGEALCKVLIGFYGNN 161
QY 197 YCSILLLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKOE 253
Db 162 YCSILFMTCLSVORYVIVNPM---GHPRKK-ANIAVGVSALWLLIFLVTIPLYVMKOT 217
QY 254 YVLVQPDITTCDDVHNTCESSPPQLY-----YFISLAPFGFLIPFVLIICYAAIIRT 307
Db 218 IYIPALNITTCDDV-----LPEEVLVGMDFNYFLSLAIGVLFPAALLTASAYVLMIKT 270
QY 308 LNA-----YDHRWLWYVKASLLIIVFTICFAPSNIILIIHHANYNNNTDGLYFIYLI 361
Db 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNLLLVVHFLIKTORQSHVYALYLV 330
QY 362 ALCLGSLNSCLDPFLYFLMSKT-RNHS 387
Db 331 ALCLSTLNSCIDPFVYFVSKDFRDHA 357

RESULT 5
US-08-476-976-61
; Sequence 61, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
```

```
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-61

Query Match      28.1%; Score 606.5; DB 2; Length 399;
Best Local Similarity 35.4%; Pred. No. 7e-42;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;

QY 22 ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPNSFEFFPSALEGWT 79
DB 5 SLAWLLGGITLLAASVSCSRTENLAPGRNNSKGRSLIGRLETOPP-----IT 51
QY 80 GATITVK--IKCPESASHLHVKNATMGVLTSSLTKLIPAIYLLVVFVGPANAVTLWM 137
DB 52 GKGVPEPGFSIDFSAS-----ILTGKLTTFVLPVVYIIVFVIGLPSNGMALWI 101
QY 138 LFRTSICITTVFY--TNLAIADELFCVTLPPKXAYHLNGNNVFGVLCRAITVIFYGNM 196
DB 102 FLFRTRKXHPAVIYMANLADLLSVIWFPLKISYHLHGNWVYGEALCKVLIGFFYGNM 161
QY 197 YCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKOE 253
DB 162 YCSILFMTCLSVQRYWVIVNPM---GHPRKK-ANIAVGVSFLAIGVFLFPALLTASAYVLMIKT 270
QY 254 YVLVQPDITTCVDVNTCESSPFQLY-----YFISLAFPGFLIPFVLIICYAAIIRT 307
DB 218 IYIPALNITTCVDV-----LPBEVLGDMFNFLSLAIGVFLFPALLTASAYVLMIKT 270
QY 308 LNA-----YDHRWLWYVKASLLIIVFTICFAPSNIILIIHHANYVYNNTDGLYFIYLI 361
DB 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNNLLLVVHYFLIKTORQSHVYALYLV 330
QY 362 ALCLGSLNSCLDPFLYFLMSKT--RNHS 387
DB 331 ALCLSTLNSCIDPFVYVFSKDFRDHA 357

RESULT 6
US-08-474-410-61
; Sequence 61, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-410-61

Query Match      28.1%; Score 606.5; DB 3; Length 399;
Best Local Similarity 35.4%; Pred. No. 7e-42;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;

QY 22 ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPNSFEFFPSALEGWT 79
DB 5 SLAWLLGGITLLAASVSCSRTENLAPGRNNSKGRSLIGRLETOPP-----IT 51
QY 80 GATITVK--IKCPESASHLHVKNATMGVLTSSLTKLIPAIYLLVVFVGPANAVTLWM 137
DB 52 GKGVPEPGFSIDFSAS-----ILTGKLTTFVLPVVYIIVFVIGLPSNGMALWI 101
QY 138 LFRTSICITTVFY--TNLAIADELFCVTLPPKXAYHLNGNNVFGVLCRAITVIFYGNM 196
DB 102 FLFRTRKXHPAVIYMANLADLLSVIWFPLKISYHLHGNWVYGEALCKVLIGFFYGNM 161
QY 197 YCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLYMLPFFILKOE 253
DB 162 YCSILFMTCLSVQRYWVIVNPM---GHPRKK-ANIAVGVSFLAIGVFLFPALLTASAYVLMIKT 270
QY 254 YVLVQPDITTCVDVNTCESSPFQLY-----YFISLAFPGFLIPFVLIICYAAIIRT 307
DB 218 IYIPALNITTCVDV-----LPBEVLGDMFNFLSLAIGVFLFPALLTASAYVLMIKT 270
QY 308 LNA-----YDHRWLWYVKASLLIIVFTICFAPSNIILIIHHANYVYNNTDGLYFIYLI 361
DB 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNNLLLVVHYFLIKTORQSHVYALYLV 330
QY 362 ALCLGSLNSCLDPFLYFLMSKT--RNHS 387
DB 331 ALCLSTLNSCIDPFVYVFSKDFRDHA 357

RESULT 7
US-08-486-673B-61
; Sequence 61, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
```

; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 61
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-486-673B-61

Query Match 28.1%; Score 606.5; DB 4; Length 399;
Best Local Similarity 35.4%; Pred. No. 7e-42;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;
QY 22 ALFAAGULLLPTFCQSGMEN--DTNNLAKTLPKTPRGAPPNSFEFFPSALEGWT 79
DB 5 SLAWLLGGITLLAASVCSRTENLAPGRNNSKGRSLIGRLETOPP-----IT 51
QY 80 GATITVK--IKPEESASHLVKNATMGVLTSSITKLIPAIYLLVGVVGPANAVTLAM 137
DB 52 GKGVPVEPGSIDFSAS-----ILTGKLTTFVLPVVIIVFVIGLPSNGMALWI 101
QY 138 LFRTSRICITVFY-TNLAIADFLFCVTLFPFKIAYHLGNWVFGVLCRATTVIIFYGNM 196
DB 102 FLRTKKKPAVIYMANLADLLSVIWFPLKISYHLGNWVYGEALCKVLIGFFYGNM 161
QY 197 YCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLVMLPFFILKQE 253
DB 162 YCSILFMTCLSVQRYWVIVNPM---GHPKKK-ANIAVGVSALAIWLLIFLVTIPLYVMKQT 217
QY 254 YVLVQPDITTCVDVHNTCESSSPQLY-----YFISLAPFGFLIPVLIYCYAALIRT 307
DB 218 IYIPALNITTCVDV-----LPEEVLVGMFNYFLSLAIGVFLPALLTASAYVLMIKT 270
QY 308 LNA-----YDRHRLWVVKASLILVITFCFAPSNIILIIHHANYNNNTDGLYFYLI 361
DB 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNNLLVHVHFLIKTORQSHVYALYLV 330
QY 362 ALCLGSLNSCLDPFLYFLMSKT-RNHS 387
DB 331 ALCLSTLNSCIDPFVYFVSKDFRDA 357

RESULT 8

US-08-097-938-2
; Sequence 2, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,938
; FILING DATE: 26-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22803-20006.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763

; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-097-938-2

Query Match 28.1%; Score 605.5; DB 1; Length 395;
Best Local Similarity 35.0%; Pred. No. 8.3e-42;
Matches 133; Conservative 70; Mismatches 112; Indels 65; Gaps 12;
QY 30 LLLLPTFCQSGMENDTNL--AKPTLPKTPKTRGAPPN---SFEPPFSALEGWTCATIT 84
DB 17 VMILLRFLCTGRNNSKGRSLIGRLETOPPITGKGVPEPGFSIDEF-----62
QY 85 VKIKPEESASHLVKNATMGVLTSSITKLIPAIYLLVGVVGPANAVTLWMLFFRTSR 144
DB 63 -----SAS-----ILTGKLTTFVLPVVIIVFVIGLPSNGMALWIFLRTKK 104
QY 145 ICTTVFY-TNLAIADFLFCVTLFPFKIAYHLGNWVFGVLCRATTVIIFYGNMYCSILL 203
DB 105 KHPAVIYMANLADLLSVIWFPLKISYHLGNWVYGEALCKVLIGFFYGNMYCSILPM 164
QY 204 ACISINRYLAIVHPFTYRGLPKHTYALVTCGL---VWATVFLVMLPFFILKQEYLYVQPD 260
DB 165 TCLSVQRYWVIVNPM---GHPKKK-ANIAVGVSALAIWLLIFLVTIPLYVMKQTIYIPALN 220
QY 261 ITTCHDVHNTCESSSPQLY-----YFISLAPFGFLIPVLIYCYAALIRTNA-----310
DB 221 ITTCHDV-----LPEEVLVGMFNYFLSLAIGVFLPALLTASAYVLMIKTFRSSAMD 273
QY 311 --YDRHRLWVVKASLILVITFCFAPSNIILIIHHANYNNNTDGLYFYLIYALCLGSL 368
DB 274 EISENKRQRAIRLIITVLAMYFICFAPSNNLLVHVHFLIKTORQSHVYALYLVCLSTL 333
QY 369 NSCLDPFLYFLMSKT-RNHS 387
DB 334 NSCIDPFVYFVSKDFRDA 353

RESULT 9

US-08-476-000-2
; Sequence 2, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-840-2

Query Match      28.1%; Score 605.5; DB 1; Length 395;
Best Local Similarity 35.0%; Pred. No. 8.3e-42;
Matches 133; Conservative 70; Mismatches 112; Indels 65; Gaps 12;

QY   30 LLLLLLTFQSGMENDTNML--AKPTLPDKTRGAPPN---SPEEPFFSALEGWTGATIT 84
Db   17 VMILLRFLCTGRNNKRSGLRGLETQPPIITCKGVVEFGFSIDEF-----62

QY   85 VKIKCEESASHLVHVKNGMTGYLTSSLSLKLIPIAYLLVVFVGVPANAVTLWMLFFRTSR 144
Db   63 -----SAS-----ILTGKLTIVFVVVIIIVFVLGPSNGHALWIFLFRTKK 104

QY   145 ICTVFY-TNLAIADFLCFTLPFKIAYHLGNMNWVFEVLCRATTVIIFYGNMYCSILL 203
Db   105 KHPAVIYMANLALDLSLVIMPLKISYHLHGNNWYVSEALCKVLIGFFYGNMYCSILPM 164

QY   204 ACISINRYLIAVHPFTYRGLPKHTYALVTCGL---VMATVELYMLPFFILKQBYYLQVDP 260
Db   165 TCLSVORYVWVNPM---GHPRKK-ANIAGVGSIAIWLLIFLVTIPLYVMKOTIYIPALN 220

QY   261 ITTCHDVNTCESSSPFOLY-----YFISLAFFGLIPFVLIIYCYAAIIRTLNA--- 310
Db   221 ITTCHDV-----LPEEVLVGMFMFYLSLAGVFLFPALLTASAYVLMIKLTRSAM 273

QY   311 --YDRHWLVWKASLLILVIFTICFAPSNIILIIHHANYYYNNNDGLYFTYLIACLGSL 368
Db   274 ERSENKRQAIRLIITVLAMYFCFAPSNLLLVVHVFLIKTQRQSHVYALYVALCLSTL 333

QY   369 NSCLDPFLYFLMSKT-RNHHS 387
Db   334 NSCIDPFVYVYSKDFRDHA 353

RESULT 11
US-08-476-976-2
; Sequence 2, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/476,976
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/390,301
;; FILING DATE: 25-JAN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ADLER, REID G.
;; REGISTRATION NUMBER: 30,988
;; REFERENCE/DOCKET NUMBER: 2803-0006.20
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 887-1500
;; TELEFAX: (202) 887-0763
;; TELEX: 90-4030
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 395 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-476-976-2

Query Match 28.1%; Score 605.5; DB 2; Length 395;
Best Local Similarity 35.0%; Pred. No. 8.3e-42;
Matches 133; Conservative 70; Mismatches 112; Indels 65; Gaps 12;
QY 30 LLLLLPTFCQSGMENDTNL--AKPTLPKTRFGAPN---SPEEPFSALEGWTCATIT 84
DB 17 VMILLRFLCTGRNNSKGRSLIGRLETPPITGKGVPEPGFSIDF-----62
QY 85 VKIKCPESASHLHVKNATMGYLTSLSSTKLIPAIYLLVFGVGPANAVTLMLFFRTRS 144
DB 63 -----SAS-----ILTGKLTTFVLPVYIIVFVIGLPSNGMALWIFLFRTKK 104
QY 145 ICTTVFY--TNLAIDFLFCVTLFPKIAHYLHNGNWNVFGVLCRAATTVIFGNYMCISILL 203
DB 105 KHPAVIYMANLADLLSVIFWFLPKISYHLHNGNWNVYGEALCKVLIGFFGNYMCISILFM 164
QY 204 ACISINRYLAIVHPFTYRGKPKHYALVTGCL--VMATVFLYMLPFILKQEVYLVQPD 260
DB 165 TCSLVQRYWVIVNPM---GHPKK-ANIAVGSLAIWLLIFLVTIPLYVWKQTIYIPALN 220
QY 261 ITTCHDVHNTCESSSPFQLY-----YFISLAFPGFLIPFVLIYCYAAIIRTLNA---310
DB 221 ITTCHDV-----LPEEVLGDMFNFLSLAIGVFLFPALLTASAYVLMKILRSSAMD 273
QY 311 --YDHRWLWYVKASLLILVITFCAPSNIILIIHANYNNYNTDGLYFIYLIACLGSL 368
DB 274 EHSNKRQRAIRLIITVLAMFYICFAPSNNLLVHVFLIKTORQSHVYALYVALCLSTL 333
QY 369 NSCLDPPELYFLMSKT--RNHS 387
DB 334 NSCIDPFVYVYVSKDFRHA 353

RESULT 12
US-08-474-410-2
; Sequence 2, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA

;; ZIP: 20006-1812
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/474,410
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/390,301
;; FILING DATE: 25-JAN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: ADLER, REID G.
;; REGISTRATION NUMBER: 30,988
;; REFERENCE/DOCKET NUMBER: 2803-0006.20
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 887-1500
;; TELEFAX: (202) 887-0763
;; TELEX: 90-4030
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 395 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-474-410-2

Query Match 28.1%; Score 605.5; DB 3; Length 395;
Best Local Similarity 35.0%; Pred. No. 8.3e-42;
Matches 133; Conservative 70; Mismatches 112; Indels 65; Gaps 12;
QY 30 LLLLLPTFCQSGMENDTNL--AKPTLPKTRFGAPN---SPEEPFSALEGWTCATIT 84
DB 17 VMILLRFLCTGRNNSKGRSLIGRLETPPITGKGVPEPGFSIDF-----62
QY 85 VKIKCPESASHLHVKNATMGYLTSLSSTKLIPAIYLLVFGVGPANAVTLMLFFRTRS 144
DB 63 -----SAS-----ILTGKLTTFVLPVYIIVFVIGLPSNGMALWIFLFRTKK 104
QY 145 ICTTVFY--TNLAIDFLFCVTLFPKIAHYLHNGNWNVFGVLCRAATTVIFGNYMCISILL 203
DB 105 KHPAVIYMANLADLLSVIFWFLPKISYHLHNGNWNVYGEALCKVLIGFFGNYMCISILFM 164
QY 204 ACISINRYLAIVHPFTYRGKPKHYALVTGCL--VMATVFLYMLPFILKQEVYLVQPD 260
DB 165 TCSLVQRYWVIVNPM---GHPKK-ANIAVGSLAIWLLIFLVTIPLYVWKQTIYIPALN 220
QY 261 ITTCHDVHNTCESSSPFQLY-----YFISLAFPGFLIPFVLIYCYAAIIRTLNA---310
DB 221 ITTCHDV-----LPEEVLGDMFNFLSLAIGVFLFPALLTASAYVLMKILRSSAMD 273
QY 311 --YDHRWLWYVKASLLILVITFCAPSNIILIIHANYNNYNTDGLYFIYLIACLGSL 368
DB 274 EHSNKRQRAIRLIITVLAMFYICFAPSNNLLVHVFLIKTORQSHVYALYVALCLSTL 333
QY 369 NSCLDPPELYFLMSKT--RNHS 387
DB 334 NSCIDPFVYVYVSKDFRHA 353

RESULT 13
US-08-937-938-5
; Sequence 5, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,938
FILING DATE: 26-JUL-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 22803-20006.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-097-938-5

Query Match 28.0%; Score 603.5; DB 1; Length 395;
Best Local Similarity 35.0%; Pred. No. 1.2e-41;
Matches 133; Conservative 69; Mismatches 113; Indels 65; Gaps 12;

QY 30 LLLLLPFCQSGMENDNNL--AKPTLPKTRGAPPN---SPEEPFSALEGTGATIT 84
DB 17 VMILLRFLCTGRNNSKGRSLGRLETPPTGKGVPEPGFSIDF----- 62
QY 85 VKIKPEESASHLHVKNATMGYLTSSLTSLKLIPIAYLLVGVGVANAVTLWMLFRTS 144
DB 63 -----SAS-----ILTGKLTTFVPPVYIIIVGLPSNGMALWIFLRTTK 104
QY 145 ICTTVFY-TNLAIAADFLFCVTLFPFKIAYHLNGNNWVFGVLCRATTVIFYGNNMYSILL 203
DB 105 KHPAVIYMANALADLLSVWFLPKISYHLGNNWVYGEALCKVLGFFGYGNNMYSILEM 164
QY 204 ACISINRYLAIVHPFTYRGLPKHTYALVTGCL---VWATVFLMPLFFILKQEYVLVQPD 260
DB 165 TCLSVQRYWVIVNPM---GHPKK-ANIAVGSVSLAIWLLIFLVTIPLYVMKQTIYIPALN 220
QY 261 ITTCHDVHNTCESSSPFOLY-----YFISLAFPGFLIPVLIYCYAAIIRTNA--- 310
DB 221 ITTCHDV-----LPEEVLVGMFNFLSLAIGVFLPALLTASAYVLMIKLTRSSAMD 273
QY 311 --YDHRWLWYKASLLILVIFTCFAPSNIILIIHHANNYYNNNTDGLYFIYLIACLGSL 368
DB 274 EHSEKKQRAIRLIITVLAMYFCFAPSNNLLLVVHYFLIKTQSQSHVYALYVALCLSTL 333
QY 369 NSCLDPFLYFLMSKT-RNHS 387
DB 334 NSCIDPFVYFVSKDFRHA 353

RESULT 14
US-08-476-000-5
Sequence 5, Application US/08476000
Patent No. 5716789
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
TITLE OF INVENTION: 63
NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,000
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-476-000-5

Query Match 28.0%; Score 603.5; DB 1; Length 395;
Best Local Similarity 35.0%; Pred. No. 1.2e-41;
Matches 133; Conservative 69; Mismatches 113; Indels 65; Gaps 12;

QY 30 LLLLLPFCQSGMENDNNL--AKPTLPKTRGAPPN---SPEEPFSALEGTGATIT 84
DB 17 VMILLRFLCTGRNNSKGRSLGRLETPPTGKGVPEPGFSIDF----- 62
QY 85 VKIKPEESASHLHVKNATMGYLTSSLTSLKLIPIAYLLVGVGVANAVTLWMLFRTS 144
DB 63 -----SAS-----ILTGKLTTFVPPVYIIIVGLPSNGMALWIFLRTTK 104
QY 145 ICTTVFY-TNLAIAADFLFCVTLFPFKIAYHLNGNNWVFGVLCRATTVIFYGNNMYSILL 203
DB 105 KHPAVIYMANALADLLSVWFLPKISYHLGNNWVYGEALCKVLGFFGYGNNMYSILEM 164
QY 204 ACISINRYLAIVHPFTYRGLPKHTYALVTGCL---VWATVFLMPLFFILKQEYVLVQPD 260
DB 165 TCLSVQRYWVIVNPM---GHPKK-ANIAVGSVSLAIWLLIFLVTIPLYVMKQTIYIPALN 220
QY 261 ITTCHDVHNTCESSSPFOLY-----YFISLAFPGFLIPVLIYCYAAIIRTNA--- 310
DB 221 ITTCHDV-----LPEEVLVGMFNFLSLAIGVFLPALLTASAYVLMIKLTRSSAMD 273
QY 311 --YDHRWLWYKASLLILVIFTCFAPSNIILIIHHANNYYNNNTDGLYFIYLIACLGSL 368
DB 274 EHSEKKQRAIRLIITVLAMYFCFAPSNNLLLVVHYFLIKTQSQSHVYALYVALCLSTL 333
QY 369 NSCLDPFLYFLMSKT-RNHS 387
DB 334 NSCIDPFVYFVSKDFRHA 353

RESULT 15
US-08-472-840-5
Sequence 5, Application US/08472840
Patent No. 5763575
GENERAL INFORMATION:


```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-000-4

Query Match      27.1%; Score 585.5; DB 1; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;

QY      80 GATITVKIKCPESASHLHVKNATMGVLTSSLSKLIPIALYLLVFGVGPANAVTLWMLF 139
DB      53 GVTVEVFSVDFESAS-----VLTKGLTTVFLPIVTVTVFVGLPSNGMALWVFL 102

QY      140 FRTSICITTVFY-TNLAIADFLFCVTLPKFIAYHLNGNNWVGEVLCRATTVIFYGNMYC 198
DB      103 FRTKKKHPAIVYMANLADLLSVIWPFLKIAYHIGNNNIYGEALCNVLIGFFYGNMYC 162

QY      199 SILLACISINRYLAIVHPFTYRGPKHTYALVTCGLVWATVFLYMLPFFILKOEYVLVQ 258
DB      163 SILFMTCVSVQRYWVIVNPMGH-SRKKANAIAGISLAIWLLLVLTPIYVVKQTIFIPA 221

QY      259 PDITTCCHDVNTCESSPPOLY-----YFISLAFGFLIPFLVLIYCYAAIIRTLNA-- 310
DB      222 LNIITTCADV-----LPEQLLVGDMFNFLSLAIGVFLPAFLTASAYVLMIRLSSA 274

QY      311 ----YDHRWLWYVKASLLILVITFCAPSNIILIIHHANYNNNDG---LYFIYLIAL 363
DB      275 MDENSEKKRKRAIKLIIVTLVAMYLICFTPSNLLLVH---YFLIKSQGSHVVALYVAL 331

QY      364 CLGSLNSCLDPPFLYFLMSKT-RNHS-TAYLTXXXNDLRE 400
DB      332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 22
US-08-472-840-4
; Sequence 4, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
```

```
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-840-4

Query Match      27.1%; Score 585.5; DB 1; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;

QY      80 GATITVKIKCPESASHLHVKNATMGVLTSSLSKLIPIALYLLVFGVGPANAVTLWMLF 139
DB      53 GVTVEVFSVDFESAS-----VLTKGLTTVFLPIVTVTVFVGLPSNGMALWVFL 102

QY      140 FRTSICITTVFY-TNLAIADFLFCVTLPKFIAYHLNGNNWVGEVLCRATTVIFYGNMYC 198
DB      103 FRTKKKHPAIVYMANLADLLSVIWPFLKIAYHIGNNNIYGEALCNVLIGFFYGNMYC 162

QY      199 SILLACISINRYLAIVHPFTYRGPKHTYALVTCGLVWATVFLYMLPFFILKOEYVLVQ 258
DB      163 SILFMTCVSVQRYWVIVNPMGH-SRKKANAIAGISLAIWLLLVLTPIYVVKQTIFIPA 221

QY      259 PDITTCCHDVNTCESSPPOLY-----YFISLAFGFLIPFLVLIYCYAAIIRTLNA-- 310
DB      222 LNIITTCADV-----LPEQLLVGDMFNFLSLAIGVFLPAFLTASAYVLMIRLSSA 274

QY      311 ----YDHRWLWYVKASLLILVITFCAPSNIILIIHHANYNNNDG---LYFIYLIAL 363
DB      275 MDENSEKKRKRAIKLIIVTLVAMYLICFTPSNLLLVH---YFLIKSQGSHVVALYVAL 331

QY      364 CLGSLNSCLDPPFLYFLMSKT-RNHS-TAYLTXXXNDLRE 400
DB      332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 23
US-08-476-976-4
; Sequence 4, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-4

Query Match 27.1%; Score 585.5; DB 2; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;

QY 80 GATITVKIKCPESASHLVKNATMGYLTSSLTKLIPAIYLLVVFVGVGPANAVTLWMLF 139
DB 53 GVTVEVFSVDFESAS-----VLTGKLTTFVLPVIVTVFVGLPSNGMALWVFL 102
QY 140 FRTRISCTTVFY-TNLAIADFLFCVTLPFKIAHYHLNGNNWVGEVLCRATTVIFYGNMYC 198
DB 103 FRTKKHHPAIVYMANLADLLSVIWFPLKIAHYHGNWNIYGEALCNVLIGFFYGNMYC 162
QY 199 SIILLACISINRYLAIVHPFTYRGPLKHYALVTCGLVWATVFLYMLPFFILKQEYVLVQ 258
DB 163 SILFMTCLSVQRVYVWNPNGH-SRKANATAGISLAIWLLILVTLPIYVVKQTIFIPA 221
QY 259 PDITTCCHDVHNTCESSSPFQLY-----YFISLAFPGFLIPFVLIICYAAIIRTINA-- 310
DB 222 LNIITTCADV-----LPEQLLVGDMFNFLSLAIGVFLPAPLTASAVLMIRLSSA 274
QY 311 ----YDHRWLWYKVASLLILVITFCAPSNIILIIHANYNNYNTDG----LYFIYLIAL 363
DB 275 MDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVH---YFLIKSQGQSHVVALYVAL 331
QY 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXXNDLRE 400
DB 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 24
US-08-474-410-4
; Sequence 4, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-410-4

Query Match 27.1%; Score 585.5; DB 3; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;

QY 80 GATITVKIKCPESASHLVKNATMGYLTSSLTKLIPAIYLLVVFVGVGPANAVTLWMLF 139
DB 53 GVTVEVFSVDFESAS-----VLTGKLTTFVLPVIVTVFVGLPSNGMALWVFL 102
QY 140 FRTRISCTTVFY-TNLAIADFLFCVTLPFKIAHYHLNGNNWVGEVLCRATTVIFYGNMYC 198
DB 103 FRTKKHHPAIVYMANLADLLSVIWFPLKIAHYHGNWNIYGEALCNVLIGFFYGNMYC 162
QY 199 SIILLACISINRYLAIVHPFTYRGPLKHYALVTCGLVWATVFLYMLPFFILKQEYVLVQ 258
DB 163 SILFMTCLSVQRVYVWNPNGH-SRKANATAGISLAIWLLILVTLPIYVVKQTIFIPA 221
QY 259 PDITTCCHDVHNTCESSSPFQLY-----YFISLAFPGFLIPFVLIICYAAIIRTINA-- 310
DB 222 LNIITTCADV-----LPEQLLVGDMFNFLSLAIGVFLPAPLTASAVLMIRLSSA 274
QY 311 ----YDHRWLWYKVASLLILVITFCAPSNIILIIHANYNNYNTDG----LYFIYLIAL 363
DB 275 MDENSEKKRKRAIKLIVTVLAMYLICFTPSNLLLVH---YFLIKSQGQSHVVALYVAL 331
QY 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXXNDLRE 400
DB 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 25
US-08-486-673B-4
; Sequence 4, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-08-486-673B-4

Query Match 27.1%; Score 585.5; DB 4; Length 398;
Best Local Similarity 36.6%; Pred. No. 3.5e-40;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;

QY 80 GATITVKIKCPESASHLVKNATMGVLTSSLTSLKLPALYLLVFFVGVGPANAVTLWMLF 139
DB 53 GVIVEIVFSDERSAS-----VLTGKLTIVFLPIVYIIVFVGVLPNGMALWFL 102

QY 140 FRTRISCTTVFY-TNLAIAFLFCVTLPFKIAIHLNGNNVFGVLCRATTIVFYGNMYC 198
DB 103 FRTKKKHPAVIYMANLADLLSVIWPPLKIAIYHIGNNNIYGEALCNVLIGFFYGNMYC 162

QY 199 SILLACISINRYLAIVHPYTYRGLPKHTYALVTCGLVWATVFLYMLPFILKQEVYLVQ 258
DB 163 SILFMTCLSVQRVYVWVNPNGH-SRKKANAIAGISLAIMLLIIVTPIVYVQTFIPA 221

QY 259 PDITTCCHDVHNTCESSPPFQLY-----YFISLAFFGFLIPFVLIIVCYAAIIRTLNA-- 310
DB 222 LNTTCHDV-----LPEQLLVGDMFNYFLSLAIGVFLPAFLTASAVVLMIRLSSA 274

QY 311 -----YDHRWLWYVKASLLIIVFTICFAPSNIILIIHANYNNYNDG---LYFIYLIAL 363
DB 275 MDENSEKKRAIKLIVTVLAMYLICFTPSNLLLVH---YFLIKSQGOSHVALYIVAL 331

QY 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXNDLRE 400
DB 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 26

US-08-486-673B-6

; Sequence 6, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Protein sequence from cDNA

US-08-486-673B-6

Query Match 27.1%; Score 584.5; DB 4; Length 398;
Best Local Similarity 36.3%; Pred. No. 4.3e-40;
Matches 123; Conservative 64; Mismatches 113; Indels 39; Gaps 10;

QY 80 GATITVKIKCPESASHLVKNATMGVLTSSLTSLKLPALYLLVFFVGVGPANAVTLWMLF 139
DB 53 GVIVEIVFSDERSAS-----VLTGKLTIVFLPIVYIIVFVGVLPNGMALWFL 102

QY 140 FRTRISCTTVFY-TNLAIAFLFCVTLPFKIAIHLNGNNVFGVLCRATTIVFYGNMYC 198
DB 103 FRTKKKHPAVIYMANLADLLSVIWPPLKIAIYHIGNNNIYGEALCNVLIGFFYGNMYC 162

QY 199 SILLACISINRYLAIVHPYTYRGLPKHTYALVTCGLVWATVFLYMLPFILKQEVYLVQ 258
DB 163 SILFMTCLSVQRVYVWVNPNGH-SRKKANAIAGISLAIMLLIIVTPIVYVQTFIPA 221

QY 259 PDITTCCHDVHNTCESSPPFQLY-----YFISLAFFGFLIPFVLIIVCYAAIIRTLNA-- 310

DB 222 LNTTCHDV-----LPEQLLVGDMFNYFLSLAIGVFLPAFLTASAVVLMIRLSSA 274

QY 311 -----YDHRWLWYVKASLLIIVFTICFAPSNIILIIHANYNNYNDG---LYFIYLIAL 363
DB 275 MDENSEKKRAIKLIVTVLAMYLICFTPSNLLLVH---YFLIKSQGOSHVALYIVAL 331

QY 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXNDLRE 400
DB 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 27

US-08-742-440A-8

; Sequence 8, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihara, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 394 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-742-440A-8

Query Match 27.0%; Score 583; DB 2; Length 394;
Best Local Similarity 35.1%; Pred. No. 5.6e-40;
Matches 136; Conservative 65; Mismatches 139; Indels 48; Gaps 13;

QY 28 AGLLLLLPTFCGSGMENDTNNAKPLTKPITKTRGAPPNSFEFFPSALSGWTCATITVKI 87
DB 12 AAIIAASLSC-SGTIGTNRSRKSLIGKVDGT-----SHVTG-KGVTVETVF 59

QY 88 KCPEBSASHLVKNATMGVLTSSLTSLKLPALYLLVFFVGVGPANAVTLWMLPRTSICT 147
DB 60 SVDEFSA-----VLTGKLTIVFLPIVYIIVFVGVLPNGMALWFLRTKKHP 109

QY 148 TVFY-TNLAIAFLFCVTLPFKIAIHLNGNNVFGVLCRATTIVFYGNMYCSILLACI 206

Db 110 AVIYMANLADLLSVTFPKTIAYHIGNNNWYGEALCNVLIGFFYGNMYCSILFWTCL 169
Qy 207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQBYLVQPDITTCDD 266
Db 170 SVQRYWVIVNPMGH-SRKKANIAIGISLAIMLLLVITPLVYVVKQTIPIPALNITTCDD 228
Qy 267 VHTTCSSSPQLYY---FISLAFPGFLIPFVLLIYCYAAIIRTLNA-----YDHRWLW 317
Db 229 V-----LPEQLLVGDPFSLAIGVFLPAPLFTASAYVLMIMLRSSAMDENSEKKRR 281
Qy 318 YVKASLLILVITFCAPSNIILIIHHANYNNYNTDG---LYFIYIALCLGSLNSCLDP 374
Db 282 AILKIVTLVAMYLICFTPSNLLLVH---YFLIKSQGQSHVAYALYVALCLSTLNSCIDP 338
Qy 375 FLYFLMSKT-RNHS-TAYLTKXXNDLRE 400
Db 339 FVYFVSHDFRDHAKNALLCRSVRTVKQ 366

RESULT 28

US-08-097-938-6
; Sequence 6, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,938
; FILING DATE: 26-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22803-20006.00
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-097-938-6

Query Match 27.0%; Score 581.5; DB 1; Length 398;
Best Local Similarity 36.0%; Pred. No. 7.5e-40;
Matches 122; Conservative 65; Mismatches 113; Indels 39; Gaps 10;
Qy 80 GATITVKIKCPESASHLVKNATMGVLTSSLTSLKLIPIAYLLVVFVGVGPANAVTLWMLF 139
Db 53 GVIVEIVFSVDFESAS-----VLTKGLTTVFLPIVYIIVFVVGPSNGMALWVFL 102
Qy 140 FRTSICITTFY-TNLAIADFLCVTLPPKIAHYLNGNNWVFGVLCRATTVIFYGNMYC 198
Db 103 FRTKKHPAVIYMANLADLLSVTFPKTIAYHIGNNNWYGEALCNVLIGFFYGNMYC 162
Qy 199 SILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQBYLVQ 258

Db 163 SILFMTCLSVQRYWVIVNPMGH-SRKKANIAIGISLAIMLLLVITPLVYVVKQTIPIA 221
Qy 259 PDITTCDDVHNTCESSSPQLYY---YFISLAFPGFLIPFVLLIYCYAAIIRTLNA-- 310
Db 222 LMITTCDDV-----LPEQVLVGMFNFVLSLAIGVFLPAPLFTASAYVLMIMLRSSA 274
Qy 311 ----YDHRWLWVVKASLLILVITFCAPSNIILIIHHANYNNYNTDG---LYFIYIAL 363
Db 275 MDENSEKKRKAUKLIVTVLAMYLICFTPSNLLLVH---YFLIKSQGQSHVAYALYVAL 331
Qy 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTKXXNDLRE 400
Db 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 29

US-08-476-000-6
; Sequence 6, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.

; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-476-000-6

Query Match 27.0%; Score 581.5; DB 1; Length 398;
Best Local Similarity 36.0%; Pred. No. 7.5e-40;
Matches 122; Conservative 65; Mismatches 113; Indels 39; Gaps 10;
Qy 80 GATITVKIKCPESASHLVKNATMGVLTSSLTSLKLIPIAYLLVVFVGVGPANAVTLWMLF 139
Db 53 GVIVEIVFSVDFESAS-----VLTKGLTTVFLPIVYIIVFVVGPSNGMALWVFL 102
Qy 140 FRTSICITTFY-TNLAIADFLCVTLPPKIAHYLNGNNWVFGVLCRATTVIFYGNMYC 198
Db 103 FRTKKHPAVIYMANLADLLSVTFPKTIAYHIGNNNWYGEALCNVLIGFFYGNMYC 162

Db 103 FRTKKKHPAVIYMANLALADLLSVIWFPLKIAIHGNNWYGEALCNVLIGFFYGNMYC 162
Qy 199 SILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQ 258
Db 163 SILFMTCLSVORYWVWPNMGH-SRKKANIAIGISLAIWLLILLVITPLVYVVKQTIFIPA 221
Qy 259 PDITTCCHDVNTCESSPPOLY-----YFISLAFFGLPFPVLLIYCYAAIIRTUNA-- 310
Db 222 LNTITTCCHDV-----LPEQVLVGD MFNYFLSLAIGVFLPAPFLTASAYVLMIRLSSA 274
Qy 311 ----YDHRMLWYVKASLLIIVITFCFAPSNIILIIHANYNNNTDG---LYFIYLIAL 363
Db 275 MDENSEKKRRAIKLIVTVLWMLYLCIFPSNLLLVH---YFLIKSOGQSHVYALIVAL 331
Qy 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXNDLRE 400
Db 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 32
US-08-474-410-6
; Sequence 6, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-474-410-6

Query Match 27.08; Score 581.5; DB 3; Length 398;
Best Local Similarity 36.08; Pred. No. 7.5e-40;
Matches 122; Conservative 65; Mismatches 113; Indels 39; Gaps 10;
Qy 80 GATITVIRIKPEESASHLVKNATMGVLTSSLTKLIPALYLVFVGVGPANAVTLWMLF 139
Db 53 GVIVEIVFSDFSAS-----VLTGKLTIVFLPIYIIVFVGVGLSNGMALVFL 102
Qy 140 FRTSICITVY-TNLAIADFLFCVTLPPFKIAVHLNGNNWVGEVLCRATTIVIFYGNMYC 198

Db 103 FRTKKKHPAVIYMANLALADLLSVIWFPLKIAIHGNNWYGEALCNVLIGFFYGNMYC 162
Qy 199 SILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQ 258
Db 163 SILFMTCLSVORYWVWPNMGH-SRKKANIAIGISLAIWLLILLVITPLVYVVKQTIFIPA 221
Qy 259 PDITTCCHDVNTCESSPPOLY-----YFISLAFFGLPFPVLLIYCYAAIIRTUNA-- 310
Db 222 LNTITTCCHDV-----LPEQVLVGD MFNYFLSLAIGVFLPAPFLTASAYVLMIRLSSA 274
Qy 311 ----YDHRMLWYVKASLLIIVITFCFAPSNIILIIHANYNNNTDG---LYFIYLIAL 363
Db 275 MDENSEKKRRAIKLIVTVLWMLYLCIFPSNLLLVH---YFLIKSOGQSHVYALIVAL 331
Qy 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXNDLRE 400
Db 332 CLSTLNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQ 370

RESULT 33
US-08-486-673B-63
; Sequence 63, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-486-673B-63

Query Match 26.8%; Score 577.5; DB 4; Length 397;
Best Local Similarity 34.4%; Pred. No. 1.6e-39;
Matches 136; Conservative 64; Mismatches 144; Indels 51; Gaps 13;
Qy 28 AGLLLLPTFCOSGMENDTNLAKPTLPKTRGAPPNSFEPPPSALEGWTGATITVKI 87
Db 12 AAILLAASLSC-SGTIQGTNRSSKGRSLIGKVDGT-----SHVTG-KGVTVTVP 59
Qy 88 KCPEESASHLVKNATMGVLTSSLTKLIPALYLVFVGVGPANAVTLWMLPRTSIC 147
Db 60 SVDEFSAS-----VLAKLTIVFLPIYIIVFVGVGLSNGMALVFLRTKKHP 109
Qy 148 TVFY-TNLAIADFLFCVTLPPFKIAVHLNGNNWVGEVLCRATTIVIFYGNMYCSILLACI 206
Db 110 AVIYMANLALADLLSVIWFPLKIAIHGNNWYGEALCNVLIGFFYGNMYCSILFMTCL 169
Qy 207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQDDITCHD 266
Db 170 SVQRYWVWPNMGH-SRKKANIAIGISLAIWLLILLVITPLVYVVKQTIFIPALNITTC 228
Qy 267 VHTNCSSSPPOLY-----YFISLAFFGLPFPVLLIYCYAAIIRTUNA-----YDHR 314
Db 229 V-----LPEQLLVGD MFNYFLSLAIGVFLPAPFLTASAYVLMIRLSSAMDENSEKK 281
Qy 315 WLWYVKASLLIIVITFCFAPSNIILIIHANYNNNTDG---LYFIYLIALCLGSLNSC 371
Db 282 RKRAIKLIVTVLWMLYLCIFPSNLLLVH---YFLIKSOGQSHVYALIVALCLSTLNSC 338
Qy 372 LDPFLYFLMSKT-RNHS-TAYLTXXNDLREOQCP 404

Db 339 IDPFVYFVSHDFRDHAKNALLCRSVRTVKQMVP 373

RESULT 34

US-08-476-000-63

; Sequence 63, Application US/08476000

; Patent No. 5716789

; GENERAL INFORMATION:

; APPLICANT: SUNDELIN, JOHAN

; APPLICANT: SCARBOROUGH, ROBERT M.

; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20006-1812

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/476,000

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/390,301

; FILING DATE: 25-JAN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: ADLER, REID G.

; REGISTRATION NUMBER: 30,988

; REFERENCE/DOCKET NUMBER: 2803-0006.20

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 887-1500

; TELEFAX: (202) 887-0763

; TELEX: 90-4030

; INFORMATION FOR SEQ ID NO: 63:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 397 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-476-000-63

Query Match 26.4%; Score 569.5; DB 1; Length 397;

Best Local Similarity 34.2%; Pred. No. 7.1e-39;

Matches 135; Conservative 64; Mismatches 145; Indels 51; Gaps 13;

Qy 28 AGLLLLPTFCQSGMENDTNNAKPTLPKTFRGAPPNSFEFFPSALEGWTGATITVKI 87

Db 12 AAIIAASLSC-SGTIOQTNRSGRSLGKVDGT-----SHVTG-KGVTETVTF 59

Qy 88 KCPEESASHLVKNATMGVLTSSLTLPALYLVFVGVGPANAVTLWMLFRTSICT 147

Db 60 SVDEFSAS-----VLAKLTTFVLPVIVTVFAVGLPSNGMALVFLFRTKKHP 109

Qy 148 TVFY-TNLAIADFLFCVTLFPFKIAYHLNNGNWFGEVLCRATTIVFYGNMYCSILLACI 206

Db 110 AVIYMANLADLLSVIWFPLKIAYHIGNWIYGEALCNVLIGFFYRNMYCSILFMTCL 169

Qy 207 SINRYLAIVHPFTYRGPLKHTYALVTCGLVWATVFLYMLPFFILKQBYLVQPDITTC 266

Db 170 SVORYWIVNPMGH-SRKKANIAIGISLAIWLLTLVTPVYVVKOTIFIPALNITTC 228

Qy 267 VHNTCESSPPOLY-----YFISLAPFGLIPVLLIYCYAAILRTLNA-----YDHR 314

Db 229 V-----LPEQLLVGMFNYFSLAIGVFLFAFLTASAYLVLMRLRSAMDENSEKK 281

Qy 315 WLWYVKASLLILVITFCFAPSNIILIIHHANYNNYNTDG---LYFIYIALCLGSLNSC 371

Db 282 RKRAIKLIVTLGMYLICFTPSNLLLVH---YFLIKSQGSHVYALVALCLSTLNSC 338

Qy 372 LDPELYFLMSKT-RNHS-TAYLTXXNDLREQQP 404

Db 339 IDPFVYFVSHDFRDHAKNALLCRSVRTVKQMVP 373

RESULT 35

US-08-472-840-63

; Sequence 63, Application US/08472840

; Patent No. 5763575

; GENERAL INFORMATION:

; APPLICANT: SUNDELIN, JOHAN

; APPLICANT: SCARBOROUGH, ROBERT M.

; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20006-1812

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/472,840

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/390,301

; FILING DATE: 25-JAN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: ADLER, REID G.

; REGISTRATION NUMBER: 30,988

; REFERENCE/DOCKET NUMBER: 2803-0006.20

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 887-1500

; TELEFAX: (202) 887-0763

; TELEX: 90-4030

; INFORMATION FOR SEQ ID NO: 63:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 397 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-472-840-63

Query Match 26.4%; Score 569.5; DB 1; Length 397;

Best Local Similarity 34.2%; Pred. No. 7.1e-39;

Matches 135; Conservative 64; Mismatches 145; Indels 51; Gaps 13;

Qy 28 AGLLLLPTFCQSGMENDTNNAKPTLPKTFRGAPPNSFEFFPSALEGWTGATITVKI 87

Db 12 AAIIAASLSC-SGTIOQTNRSGRSLGKVDGT-----SHVTG-KGVTETVTF 59

Qy 88 KCPEESASHLVKNATMGVLTSSLTLPALYLVFVGVGPANAVTLWMLFRTSICT 147

Db 60 SVDEFSAS-----VLAKLTTFVLPVIVTVFAVGLPSNGMALVFLFRTKKHP 109

Qy 148 TVFY-TNLAIADFLFCVTLFPFKIAYHLNNGNWFGEVLCRATTIVFYGNMYCSILLACI 206

Db 110 AVIYMANLADLLSVIWFPLKIAYHIGNWIYGEALCNVLIGFFYRNMYCSILFMTCL 169

Qy 207 SINRYLAIVHPFTYRGPLKHTYALVTCGLVWATVFLYMLPFFILKQBYLVQPDITTC 266

Db 170 SVORYWIVNPMGH-SRKKANIAIGISLAIWLLTLVTPVYVVKOTIFIPALNITTC 228

QY 267 VHTCSESSPPOLY-----YFISLAPFGFLIPFVLIIYCYAAIIRTUNA-----YDHR 314
 Db 229 V-----LPEQLLVGDMFNYFLSLAIGVFLFAFLTASAYVLMIRMLRSSAMDESEK 281
 QY 315 WLWYKASLLILVIFTCFAPSNIILIIHANYVYNTDG--LYFIYLIALCLGSLNSC 371
 Db 282 RKRAIKLIVTVLGMVLCFTPSNLLVH--YFLIKSQSQSHVYALYIIVALCLSLNSC 338
 QY 372 LDPFLYFLMSKT-RNHS-TAYLTXXNDLREOQOP 404
 Db 339 IDPFVYVYVSHDRDHAKNALLCRSVRTVKMQVP 373

RESULT 36
 US-08-476-976-63
 ; Sequence 63, Application US/08476976
 ; Patent No. 5874400
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA: US/08/476,976
 ; APPLICATION NUMBER: 07-JUN-1995
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-0763
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 63:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 397 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-476-976-63

Query Match 26.4%; Score 569.5; DB 2; Length 397;
 Best Local Similarity 34.2%; Pred. No. 7.1e-39;
 Matches 135; Conservative 64; Mismatches 145; Indels 51; Gaps 13;

QY 28 AGLILLPTFCQSGMENDTNLAKPTLPKTFRGAPPNSFEPPFSALEGWTCATITVKI 87
 Db 12 AAILLAASLSC-SGTIGQTRSSKGRSLICKVDGT-----SHVTG-KGVTVETVF 59
 QY 88 KCPEESASHLVKVNATMGYLTSSLSKLIPIAYLLVFWGVGPANAVTLMWLFRTSIC 147
 Db 60 SVDEFSSA-----VLAGKLTIVFLPIVTVIVFAVGLPSNGMALVFLFRTKKHP 109
 QY 148 TVFY-TNLATDFLCVTLFPKTAHUNGNNWFGVLCRATTVIFGNNYCSILLIACI 206
 Db 110 AVIYMANLALADLLSVIFWFLPKIAYHIGNNWIYGEALCNVLIGFFFYRNMYCSILFMTCL 169

QY 207 SINRYLAIYVHPYTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQPDITICH 266
 Db 170 SVORYWVIVNPMGH-SRKKANIAIGISLAIMLLTLVITPLVYVKQTIPIPALNITICH 228
 QY 267 VHTCSESSPPOLY-----YFISLAPFGFLIPFVLIIYCYAAIIRTUNA-----YDHR 314
 Db 229 V-----LPEQLLVGDMFNYFLSLAIGVFLFAFLTASAYVLMIRMLRSSAMDESEK 281
 QY 315 WLWYKASLLILVIFTCFAPSNIILIIHANYVYNTDG--LYFIYLIALCLGSLNSC 371
 Db 282 RKRAIKLIVTVLGMVLCFTPSNLLVH--YFLIKSQSQSHVYALYIIVALCLSLNSC 338
 QY 372 LDPFLYFLMSKT-RNHS-TAYLTXXNDLREOQOP 404
 Db 339 IDPFVYVYVSHDRDHAKNALLCRSVRTVKMQVP 373

RESULT 37
 US-08-474-410-63
 ; Sequence 63, Application US/08474410
 ; Patent No. 6043212
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA: US/08/474,410
 ; APPLICATION NUMBER: 25-JAN-1995
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 63:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 397 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-474-410-63

Query Match 26.4%; Score 569.5; DB 3; Length 397;
 Best Local Similarity 34.2%; Pred. No. 7.1e-39;
 Matches 135; Conservative 64; Mismatches 145; Indels 51; Gaps 13;

QY 28 AGLILLPTFCQSGMENDTNLAKPTLPKTFRGAPPNSFEPPFSALEGWTCATITVKI 87
 Db 12 AAILLAASLSC-SGTIGQTRSSKGRSLICKVDGT-----SHVTG-KGVTVETVF 59
 QY 88 KCPEESASHLVKVNATMGYLTSSLSKLIPIAYLLVFWGVGPANAVTLMWLFRTSIC 147

Db 60 SVDFSGAS-----VLAKLTITVFLPIVITVFAVGLPSNGMALVFLFRTRKKHP 109
QY 148 TVFY-TNLATADFLFCVTLFPFKIAYHLNGNNWVGEVLCRATTIVIFGNNMYSIILLACI 206
Db 110 AVIYMANLADLISVIFPKIAYHGNWYIYGEALCNVLIGFFYRNNMYSILFMTCL 169
QY 207 SINRYLAIPHFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQPDITTCCHD 266
Db 170 SVQRYVIVNPMGH-SRKKANIALIGISLAIWLTLTLLVPLVYVKOTIFIPALNITTCCHD 228
QY 267 VHTNCSSSPQLY-----YFISLAPFFGLIPFVLIYYCYAAIIRTLNA-----YDHR 314
Db 229 V-----LPEQLLVGMFNYFLSLAIGVFLFPAFLTASAYVLMIIRSSAMENSEKK 281
QY 315 WLVVVKASLLIVFTICFAPSNIILIIHHANYNNYNTDG---LYFTYLIALCLGSLNSC 371
Db 282 RKRAIKLIVTLVGLMYLICFTFPSLLLVH---YFLIKSQOSHVYALYIVALCLSTLNSC 338
QY 372 LDPLFLYPLMSKT-RNHS-TAYLTXXXNDLREQQP 404
Db 339 IDPFVYFVSHDFRDHAKNALLCRSVRTVKQMOPV 373

RESULT 38

US-08-097-938-7
; Sequence 7, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,938
; FILING DATE: 26-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22803-20006.00
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-097-938-7

Query Match 24.4%; Score 526; DB 1; Length 425;
Best Local Similarity 32.9%; Pred. No. 2.6e-35;
Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;
QY 14 GTQVKKALIFAAGLILLLLPFCQSGMENDTNLAKPTLPKTRGAPPN-SPEEFPF 72
Db 2 GPRLLLVAAFCFLSCGLFLLSARTARRPESKATN-----ATLDRSLRNPNDKPEFFWE 57
QY 73 SALEGWTGAT----ITVKIKCPESASHLVKNATMGYLTSSLSLTKLIPAILLVFVVG 128

Db 58 DEEKESGLTEVRLVINSKSSPLQKQPAFISEDASGYLTSSWLTLPVPSVVTGVFWSL 117
QY 129 PAN--AVTLWMLPFRTRISCTIVFTYNLAIAADFLFCVTLFPFKIAYHLNGNNWVGEVLCR 186
Db 118 PLNIMAVIVFILMKVKK-PAVVYMLHLATADVFLSVLPFKISYFYSGDQFGSELGR 176
QY 187 ATTIVIFYGNNMYSIILLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLY 243
Db 177 FTAAFYCMYASILLMTVISIDRFNAVYPMQISWRTLGRASP---TCLAIWALAIAAG 233
QY 244 MLPFFILKQEYLVQPDITTCCHDVHTNCSSSPFQLYYFISLAPFFGLIPFVLIYYCYAA 303
Db 234 VVPLVLKEQTIQVPGNITTCCHDVLENTLEG-VYAYYFSAFSAVFFFPVPLIISTVCYVS 292
QY 304 IIRTLN-----AYDRHLWTVVKASLLILVITFCFAPSNIILIIHHANY-YYNNTDGL 355
Db 293 IIRCLSSSAVANRKSRAIF---LSAAVFCIFIIICFGPTNVLLIAHYSFLSHTSTTEAA 349
QY 356 YFIYLIALCLGSLNSCLDPFLYFLMS 381
Db 350 YFAYLLVCVSSISSCIDPLIYYAS 375

RESULT 39

US-08-476-000-7
; Sequence 7, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-476-000-7

Query Match 24.4%; Score 526; DB 1; Length 425;
Best Local Similarity 32.9%; Pred. No. 2.6e-35;
Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;

```

QY 14 GTQVIMKALIPAAAGLLLLPTFCOSGMENDTNLAKPTLPKIKTRGAPPN-SPEEFPF 72
Db 2 GPRRLILVAACFLGCLFLLSARTRARRPESKATN----ATLDRSFLLRNPNDKYEPFWE 57
QY 73 SALEGTGAT----ITVKIKCPESASHLVKNATMGYLTSSLSKLIIPAIYLLVFVGV 128
Db 58 DEKNESGLTEYRLVINSKSSPQKOLPAFISDASGYLTSSWLTFLVPSVYTGCVFVSL 117
QY 129 PAN--AVTLMMLFFRTRISCTTVFYTNLAIADELFVCTLPFKIAYHLNGNNWVFGVLCR 186
Db 118 PLNIMAIVVFIKMKVKK--PAVVYMLHLATADVLSVLPFKISYFSGDMQFGSELCL 176
QY 187 ATTIVIFGMYCISILLACISINRYLAIVHP---FTYRGLPKHYALVTCGLVWATVFLY 243
Db 177 FVTAAFYCNMYAISILLMTVISIDRFLAVVPMQSLSWRTLGRASF---TCLAIWALAIA 233
QY 244 MLPFFILKOEYVLVQPDITTCCHDVHNTCESSPFQLYYFISLAFPGFLIPFVLIICYAA 303
Db 234 VPLVLKEQTIQVGLNITTCCHDVNLNETLLEG--YAYYFSAFSAVFFVPLIISTVCYVS 292
QY 304 IIRTLN-----AYDHRWLWYVKASLLILVITFCFAPSNIILIIHHANY--YNNNTDGL 355
Db 293 IIRCLSSAVANRSKSRALF---LSAAVFCIFICFGPTNVLIIAHYFSLSHTSTTEA 349
QY 356 YFIYLIACLGSLNSCLDPFLYPLMS 381
Db 350 YFAYLLCVCVSSISSCIDPLIYYAS 375

```

RESULT 40

```

US-08-472-840-7
; Sequence 7, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-840-7

```

```

Query Match 24.4%; Score 526; DB 1; Length 425;
Best Local Similarity 32.9%; Pred. No. 2.6e-35;
Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;
QY 14 GTQVIMKALIPAAAGLLLLPTFCOSGMENDTNLAKPTLPKIKTRGAPPN-SPEEFPF 72
Db 2 GPRRLILVAACFLGCLFLLSARTRARRPESKATN----ATLDRSFLLRNPNDKYEPFWE 57
QY 73 SALEGTGAT----ITVKIKCPESASHLVKNATMGYLTSSLSKLIIPAIYLLVFVGV 128
Db 58 DEKNESGLTEYRLVINSKSSPQKOLPAFISDASGYLTSSWLTFLVPSVYTGCVFVSL 117
QY 129 PAN--AVTLMMLFFRTRISCTTVFYTNLAIADELFVCTLPFKIAYHLNGNNWVFGVLCR 186
Db 118 PLNIMAIVVFIKMKVKK--PAVVYMLHLATADVLSVLPFKISYFSGDMQFGSELCL 176
QY 187 ATTIVIFGMYCISILLACISINRYLAIVHP---FTYRGLPKHYALVTCGLVWATVFLY 243
Db 177 FVTAAFYCNMYAISILLMTVISIDRFLAVVPMQSLSWRTLGRASF---TCLAIWALAIA 233
QY 244 MLPFFILKOEYVLVQPDITTCCHDVHNTCESSPFQLYYFISLAFPGFLIPFVLIICYAA 303
Db 234 VPLVLKEQTIQVGLNITTCCHDVNLNETLLEG--YAYYFSAFSAVFFVPLIISTVCYVS 292
QY 304 IIRTLN-----AYDHRWLWYVKASLLILVITFCFAPSNIILIIHHANY--YNNNTDGL 355
Db 293 IIRCLSSAVANRSKSRALF---LSAAVFCIFICFGPTNVLIIAHYFSLSHTSTTEA 349
QY 356 YFIYLIACLGSLNSCLDPFLYPLMS 381
Db 350 YFAYLLCVCVSSISSCIDPLIYYAS 375

```

Search completed: June 24, 2003, 12:07:21
Job time : 21.0221 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:52:41 ; Search time 52.0638 Seconds
(without alignments)
1044.224 Million cell updates/sec

Title: US-09-208-629F-6

Perfect score: 2157

Sequence: 1 CSMILQISXRLRDGTQVIRM.....AVLTXXNDLREQGPSORT 408

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries.

Database :

A Genesecq_101002.*
1: /SID22/gcgdata/genesecq/genesecp-emb1/AA1980.DAT.*
2: /SID22/gcgdata/genesecq/genesecp-emb1/AA1981.DAT.*
3: /SID22/gcgdata/genesecq/genesecp-emb1/AA1982.DAT.*
4: /SID22/gcgdata/genesecq/genesecp-emb1/AA1983.DAT.*
5: /SID22/gcgdata/genesecq/genesecp-emb1/AA1984.DAT.*
6: /SID22/gcgdata/genesecq/genesecp-emb1/AA1985.DAT.*
7: /SID22/gcgdata/genesecq/genesecp-emb1/AA1986.DAT.*
8: /SID22/gcgdata/genesecq/genesecp-emb1/AA1987.DAT.*
9: /SID22/gcgdata/genesecq/genesecp-emb1/AA1988.DAT.*
10: /SID22/gcgdata/genesecq/genesecp-emb1/AA1989.DAT.*
11: /SID22/gcgdata/genesecq/genesecp-emb1/AA1990.DAT.*
12: /SID22/gcgdata/genesecq/genesecp-emb1/AA1991.DAT.*
13: /SID22/gcgdata/genesecq/genesecp-emb1/AA1992.DAT.*
14: /SID22/gcgdata/genesecq/genesecp-emb1/AA1993.DAT.*
15: /SID22/gcgdata/genesecq/genesecp-emb1/AA1994.DAT.*
16: /SID22/gcgdata/genesecq/genesecp-emb1/AA1995.DAT.*
17: /SID22/gcgdata/genesecq/genesecp-emb1/AA1996.DAT.*
18: /SID22/gcgdata/genesecq/genesecp-emb1/AA1997.DAT.*
19: /SID22/gcgdata/genesecq/genesecp-emb1/AA1998.DAT.*
20: /SID22/gcgdata/genesecq/genesecp-emb1/AA1999.DAT.*
21: /SID22/gcgdata/genesecq/genesecp-emb1/AA2000.DAT.*
22: /SID22/gcgdata/genesecq/genesecp-emb1/AA2001.DAT.*
23: /SID22/gcgdata/genesecq/genesecp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES.

Result No.	Score	Query Match	Length	ID	Description
1	1998	92.6	374	19 AAW51406	Human protease-act
2	1435.5	66.1	369	19 AAW51405	Mouse protease-act
3	606.5	28.1	399	16 AAR66922	Murine C140 recept
4	606.5	28.1	399	17 AAW01954	Murine C140 recept
5	600.5	27.8	395	16 AAR66920	Murine C140 recept
6	600.5	27.8	395	17 AAW01952	Murine C140 recept
7	589.5	27.3	397	21 AAB35641	Human PAR-2 protei
8	585.5	27.1	398	16 AAR66921	Human C140 recept
9	585.5	27.1	398	17 AAW01953	Human C140 recept
10	583	27.0	394	19 AAW51408	Human protease-act

11	569.5	26.4	397	16 AAR66923	Human C140 recept
12	569.5	26.4	397	17 AAW01955	Human C140 recept
13	531.5	24.6	396	20 AAY15081	Mouse protease-act
14	525	24.3	408	20 AAY50137	FLAG epitope-tagge
15	523.5	24.3	385	20 AAY50139	Mutant human prote
16	523	24.2	425	13 AAR27240	Human thrombin rec
17	523	24.2	425	19 AAW51407	Human protease-act
18	523	24.2	425	20 AAY49570	Human thrombin rec
19	523	24.2	425	23 AAG80697	Human thrombin pro
20	523	24.2	425	23 AAE17032	Human thrombin pro
21	523	24.2	426	21 AAY45035	Human thrombin rec
22	521.5	24.2	385	20 AAY15082	Human protease-act
23	521.5	24.2	385	20 AAY50135	Human protease-act
24	521.5	24.2	385	21 AAY45036	Human protease-act
25	521.5	24.2	385	22 AAB47623	Human PAR4. Homo
26	521	24.2	385	20 AAY50138	Mutant human prote
27	519	24.1	425	23 AAG35300	Human PAR1 type th
28	516.5	23.9	425	15 AAR60698	Fragment of the hu
29	512	23.7	892	18 AAW16314	Human thrombin rec
30	507.5	23.5	402	23 AAG35298	Human PAR1 type th
31	499.5	23.2	371	23 AAG35299	Human PAR1 type th
32	471.5	21.9	359	19 AAG69598	Mouse G-protein co
33	461	21.4	359	21 AAY71304	Human orphan G pro
34	461	21.4	359	21 AAB02838	Human G-protein co
35	460	21.3	359	21 AAY84815	A human G-protein
36	460	21.3	359	21 AAY69485	Amino acid sequenc
37	460	21.3	359	22 AAG78530	G-protein coupled
38	460	21.3	359	22 AAG80966	Human NGPCRS #2.
39	460	21.3	359	22 AAB62285	Human G-protein co
40	432.5	20.1	374	22 AAE04390	Turkey P2Y nucleot
41	428	19.8	537	23 AAY074538	Human P2Y purinoco
42	413.5	19.2	339	21 AAY90652	Human mutant G pro
43	411	19.1	322	15 AAR48715	G-protein coupled
44	411	19.1	322	17 AAW02687	G-protein coupled
45	410.5	19.0	339	15 AAR53752	Seven transmembran

ALIGNMENTS

RESULT 1
AAW51406
ID AAW51406 standard; Protein; 374 AA.
XX
AC AAW51406;
XX
DT 12-OCT-1998 (first entry)
XX
DE Human protease-activated receptor 3 (PAR3).
XX
KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
KW G-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis.
XX
OS Homo sapiens.
XX
FH Key
FT Modified-site 25..27 /note= "Asn is N-glycosylated"
FT Cleavage-site 38..39 /note= "thrombin cleavage site"
FT Peptide 48..51 /note= "hirudin-like sequence"
FT Modified-site 82..84 /note= "Asn is N-glycosylated"
FT Domain 95..117 /label= TM1
FT /note= "transmembrane domain 1"
FT Domain 126..149 /label= TM2
FT /note= "transmembrane domain 2"

FT Domain 168..191
 FT /label= TM3
 FT /note= "transmembrane domain 3"
 FT Domain 207..231
 FT /label= TM4
 FT /note= "transmembrane domain 4"
 FT Domain 261..286
 FT /label= TM5
 FT /note= "transmembrane domain 5"
 FT Domain 301..323
 FT /label= TM6
 FT /note= "transmembrane domain 6"
 FT Modified-site 331..333
 FT /note= "Asn is N-glycosylated"
 FT Domain 335..360
 FT /label= TM7
 FT /note= "transmembrane domain 7"
 FT XX
 PN WO9818456-A1.
 XX
 PD 07-MAY-1998.
 XX
 PF 29-OCT-1997; 97WO-US19732.
 XX
 PR 30-OCT-1996; 96US-0742440.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Connolly A, Coughlin SR, Ishihara H;
 XX
 DR WPI; 1998-271905/24.
 DR N-PSDB; AAV07374.
 XX
 PT DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.
 XX
 PS Claim 3; Page 41-42; 74pp; English.
 XX
 CC This polypeptide comprises human protease-activated receptor 3
 CC (PAR3), a cell surface protein which is specifically activated by
 CC thrombin or a thrombin agonist, thereby activating signalling
 CC events such as phosphoinositide hydrolysis, calcium ion efflux and
 CC platelet aggregation. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAV07372), and shows homology to human PAR1
 CC and PAR2 (see AAW51407 and AAW51408). The mouse PAR3 amino acid
 CC sequence (see AAW51405) is also provided. Also claimed are vectors,
 CC host cells and an assay device. Host cells are used to screen
 CC the effects of thrombin-PAR3 interaction. Agonists are used to
 CC treat wounds, thrombosis, atherosclerosis, restenosis, inflammation
 CC and other thrombin activated disorders. Antagonists (see AAW51415-21)
 CC are used to control blood coagulation and thereby to treat heart
 CC attack and stroke. They also mediate inflammatory and proliferative
 CC responses to injury as occur in wound healing, atherosclerosis,
 CC restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis.
 XX
 SQ Sequence 374 AA;
 Query Match 92.6%; Score 1998; DB 19; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.3e-198;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 MKALIFAAAGLLLLPTFCOSGMENDNNLAKETPLPKTRGAPPNSFEFFPSSALEGWT 79
 Db 1 MKALIFAAAGLLLLPTFCOSGMENDNNLAKETPLPKTRGAPPNSFEFFPSSALEGWT 60
 QY 80 GATITVKIKCPESASHLVKNAKTMGYLTSSLSKLIPIAYLLVFFVGVGPANAVTLWMLF 139
 Db 61 GATITVKIKCPESASHLVKNAKTMGYLTSSLSKLIPIAYLLVFFVGVGPANAVTLWMLF 120
 QY 140 FRTRISCTTVFYNTLAIADFLFCVTLPFKIAYHLNGNNWVFGVLCRAITTVIFYGNMYS 199

Db 121 FRTRISCTTVFYNTLAIADFLFCVTLPFKIAYHLNGNNWVFGVLCRAITTVIFYGNMYS 180
 QY 200 ILLACISINRYLAIVHPETRYGLPKHTYALATCGLVWATVFLYMLPFFILKQEYVLP 259
 Db 181 ILLACISINRYLAIVHPETRYGLPKHTYALATCGLVWATVFLYMLPFFILKQEYVLP 240
 QY 260 DITTCDDVHNTCESSPPFOLYYFISLAPFGFLIPFLVLIYCYAAIIRTINAYDHRWLWV 319
 Db 241 DITTCDDVHNTCESSPPFOLYYFISLAPFGFLIPFLVLIYCYAAIIRTINAYDHRWLWV 300
 QY 320 KASLLILVITTCFAPSNIILIIHHANYNNNTDGLYFIYLIYALCLGSLNSCLDPFLYFL 379
 Db 301 KASLLILVITTCFAPSNIILIIHHANYNNNTDGLYFIYLIYALCLGSLNSCLDPFLYFL 360
 QY 380 MSKTRNHSTAYLTK 393
 Db 361 MSKTRNHSTAYLTK 374
 RESULT 2
 AAW51405
 ID AAW51405 standard; Protein; 369, AA.
 XX
 AC AAW51405;
 XX
 DT 12-OCT-1998 (first entry)
 XX
 DE Mouse protease-activated receptor 3 (PAR3).
 XX
 KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
 KW G-protein coupled receptor; agonist; antagonist; thrombosis;
 KW atherosclerosis; restenosis; inflammation; blood coagulation;
 KW blood clotting; heart attack; stroke; wound healing;
 KW adult respiratory distress syndrome; glomerulosclerosis.
 XX
 OS Mus sp.
 XX
 PN WO9818456-A1.
 XX
 PD 07-MAY-1998.
 XX
 PF 29-OCT-1997; 97WO-US19732.
 XX
 PR 30-OCT-1996; 96US-0742440.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Connolly A, Coughlin SR, Ishihara H;
 XX
 DR WPI; 1998-271905/24.
 DR N-PSDB; AAV07372.
 XX
 PT DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.
 XX
 PS Claim 3; Page 39-40; 74pp; English.
 XX
 CC This polypeptide comprises mouse protease-activated receptor 3
 CC (PAR3), a cell surface protein which is specifically activated by
 CC thrombin or a thrombin agonist, thereby activating signalling
 CC events such as phosphoinositide hydrolysis, calcium ion efflux and
 CC platelet aggregation. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAV07372). The human PAR3 amino acid
 CC sequence (see AAW51406) is also provided. Also claimed are vectors,
 CC host cells and an assay device. Host cells are used to screen
 CC the effects of thrombin-PAR3 interaction. Agonists are used to
 CC treat wounds, thrombosis, atherosclerosis, restenosis, inflammation
 CC and other thrombin activated disorders. Antagonists (see AAW51415-21)
 CC are used to control blood coagulation and thereby to treat heart
 CC attack and stroke. They also mediate inflammatory and proliferative
 CC responses to injury as occur in wound healing, atherosclerosis,
 CC

CC restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis.
XX Sequence 369 AA;
SQ
Query Match 66.1%; Score 1425.5; DB 19; Length 369;
Best Local Similarity 71.6%; Pred. No. 3e-139;
Matches 260; Conservative 44; Mismatches 58; Indels 1; Gaps 1;
QY 20 MKALIFAAAGLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPNSFEFFPFSALEGWT 79
DB 1 MKLILVAAGLLFLPVTVCQSGI-NVSDNSAKPTLTIKSPNGGPQNTFEFFPLSDIEGWT 59
QY 80 GAITIVKIKCPESASHLVKNATMGVLTSSSTKLIPAIYLLVGVVPANAVTLWMLF 139
DB 60 GATTTTCAECPEBSISTLHVNNATIGYLRSSLSQVIPAIIYLLVGVVPSNVTTLWKL 119
QY 140 FRTRISCTTVFYTNLAIDFLFCVTLFPFKIAYHLGNWVFGVLCRATTIVIFYGNMYS 199
DB 120 LRTKSLVIFHNLAIDFLFCVTLFPFKIAYHLGNWVFGVLCRATTIVIFYGNMYS 179
QY 200 ILLACISINRYLAIVHPFTYRGVLTGVLVATVTCGLVWATVFLYMLPFFFLKQEYLVQP 259
DB 180 ILILTGMGINRYLATAHPFTYQKLPKESFSLMCGIVWVWVFLYMLPFFVLKQEXHLVHS 239
QY 260 DITTDHVNHTCSSSPFQLYYFISLAPFGFLIPFVLIIVCYAAIIRTNAYDHRMLWYV 319
DB 240 EITTDHVDVACESPSSEFYFVSLAFFGLIPFVIIIFCYTTLIHLKSKDRIWLGYI 299
QY 320 KASILLIVITICFAPSNIILIIHHANYNNYNTDGLYFYLIALLCLGSLNSCLDPLFLY 379
DB 300 KAVLLILVITICFAPTNIILIIHHANYNNYNTDGLYFYLIALLCLGSLNSCLDPLFLY 359
QY 380 MSK 382
DB 360 MSK 362
RESULT 3
AAW01954
ID AAW01954 standard; Protein; 399 AA.
XX
AC AAW01954;
XX
XX 22-AUG-1995 (first entry)
XX Murine C140 receptor deduced from cDNA.
XX G-protein-coupled receptor; G-protein; C140 receptor.
XX Mus musculus.
XX WO9503318-A.
XX 02-FEB-1995.
XX 26-JUL-1994; 94WO-US08536.
XX 26-JUL-1993; 93US-0097938.
XX (CORT-) COR THERAPEUTICS.
XX Scarborough RM, Sundelin J;
XX WPI; 1995-075182/10.
XX N-PSDB; AAQ84559.
XX New DNA encoding recombinant C140 receptor - and novel agonists
PT and antagonists and specific antibodies with therapeutic and
PT diagnostic applications.
XX Example; Fig 10; 57pp; English.
XX A cDNA library from a mouse stomach was constructed in lambda gt10

CC and screened with a probe encompassing the C140 genomic clone (see
CC AAQ84557). A single phage clone was isolated and cut with EcoRI. The
CC insert was cloned in pBluescript and pS5 and sequenced. The
CC complete nt sequence and deduced AA sequence is given in AAQ84559 &
CC AAR66922. 5' RACE resulted in the addition of only 27 bps to the 5'
CC end. The 5' end of the apparent coding region differs from the 5' end of
CC the cDNA sequence is correct.
XX
SQ Sequence 399 AA;
Query Match 28.1%; Score 606.5; DB 16; Length 399;
Best Local Similarity 35.4%; Pred. No. 3.1e-54;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;
QY 22 ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPNSFEFFPFSALEGWT 79
DB 5 SLAWLLGGITLLAASVSCRTENLAPGRNNSKRSGLIGRLETQPP-----IT 51
QY 80 GAITIVK--IKCPESASHLVKNATMGVLTSSSTKLIPAIYLLVGVVPANAVTLWML 137
DB 52 KGKVPVPEPGFSIDFSAS-----LITGKLTTFVLPVVVYIIVFVIGLPSNGMALWI 101
QY 138 LFFRTRISCTTVFY-TNLAIDFLFCVTLFPFKIAYHLGNWVFGVLCRATTIVIFYGNM 196
DB 102 FLFRTRKTKHPAVIYMANLALADLLSVIWPFLKISYHLGNWVYGEALCKVLIGFFYGNM 161
QY 197 YCSILLACISINRYLAIVHPFTYRGVLTGVLVATVTCGL--VWATVFLYMLPFFFLKOE 253
DB 162 YCSILFMTCLSVQRYWVIVNPM---GHPKK-ANIAVGVSALWLLIFLVTIPLYVMKQT 217
QY 254 YVLVQPDITTDHVNHTCSSSPFQLY-----YFISLAFPGFLIPFVLIIVCYAAIIR 307
DB 218 IYIPALNITTDHV-----LPEEVLVGMENYFLSLAIGVLPFALLTASAYVLMWT 270
QY 308 LNA-----YDHRWLMYVKASLLILVITICFAPSNIILIIHHANYNNYNTDGLYFYLI 361
DB 271 LRSSAMDEHSEKKRQRAIRLIITVLAMYFICFAPSNNLLVHVHFLIKTQRQSHVYALYV 330
QY 362 ALCLGSLNSCLDPLFLYFLMSKT-RNHS 387
DB 331 ALCLSLNSCLDPLFYFVSKDFRDHA 357
RESULT 4
AAW01954
ID AAW01954 standard; Protein; 399 AA.
XX
AC AAW01954;
XX
XX 02-APR-1997 (first entry)
XX Murine C140 receptor.
XX C140 receptor; G-protein linked; coupled; seven pass; agonist;
XX antagonist; hypertension; hypotension; blood pressure.
XX Mus sp.
XX
XX Key Location/Qualifiers
FH Peptide 1..31
FT /note= "the signal peptide differs from that
FT encoded by a genomic DNA sequence for
FT this receptor (see AAW01952), the signal
FT sequence given here is believed to be
FT the correct sequence"
FT Protein 32..399
FT /note= "mature protein"
XX WO9623225-A1.
XX 01-AUG-1996.
XX

Db 17 VMLLRLCTGRNNSKGRSLIGRLTQPTTGKGVPEPGFSIDEF----- 62

Qy 85 VKIKPEESASHLHVKNATMGYLTSLTKLIPAIYLLVGVGVPANAVTLMLFFRTRS 144

Db 63 -----SAS-----ILTGKLTTFVLPVWYIIVFVIGLPSNGMALWFLFRTKK 104

Qy 145 ICTTVFY-TNLAIADFLFCVTLFPKIAHYLNGNWNVFGVLCRATTVIFGNYMYSILL 203

Db 105 KHPAVIYMANLALADLLSVIFWFLKISYHLGNNWYVGEALCKVLIGFFYGNMYSILFM 164

Qy 204 ACISINRYLAIVHPFTYRGPKHYALVTCGL---VMATVFLYMLPFILKQEVYLVQPD 260

Db 165 TCLSVQRYWVIVNPM---GHPKPK-ANIAVGSLAIWLLFLVTIPDYVWKQIYIPALN 220

Qy 261 ITTCHDVHNTCESSPPQLY-----YFISLAFPGFLIPFVLIYCYAAIIRTLNA----- 310

Db 221 ITTCHDV-----LPEEVLVGMFNYFLSLAIGVFLFPALLTASAVVLMIKTLRSSAMD 273

Qy 311 --YDHRWLWYVKASLLILVITTCFAPSNIILIIHHANYNNYNTDGLFYLIYALCLGSL 368

Db 274 HSENKQRAIRLIITVLMYFICFRPSNLLLVVHYFLIKTQROSHVYALVLCISLTL 333

Qy 369 NSCLDPFLYFLMSKT-RNHS 387

Db 334 NSCIDPFVYFVSKDFRDHA 353

RESULT 6

AAW01952

ID AAW01952 standard; Protein; 395 AA.

AC AAW01952;

XX

DT 01-APR-1997 (first entry)

XX

DE Murine C140 receptor, including putative signal sequence.

XX

KW C140 receptor; G-protein linked; coupled; seven pass; agonist;

KW antagonist; hypertension; hypotension; blood pressure.

XX

OS Mus sp.

XX

Key Location/Qualifiers

FT Peptide 1..27

FT /note= "putative signal peptide, differs from

FT this receptor (see AAW01954), the signal

FT sequence given for the cDNA clone is

FT believed to be the correct sequence"

FT

FT Protein 28..395

FT /note= "mature protein"

FT

FT Modified-site 29

FT /note= "potential Asn-linked glycosylation site"

FT

FT Cleavage-site 34..35

FT /note= "putative protease receptor cleavage site"

FT Region 78..100

FT /note= "transmembrane region I"

FT Region 108..128

FT /note= "transmembrane region II"

FT Region 148..169

FT /note= "transmembrane region III"

FT Region 188..210

FT /note= "transmembrane region IV"

FT Modified-site 220

FT /note= "potential Asn-linked glycosylation site"

FT Region 244..264

FT /note= "transmembrane region V"

FT Region 286..306

FT /note= "transmembrane region VI"

FT Region 324..345

FT /note= "transmembrane region VII"

XX

PN W09623225-A1.

XX 01-AUG-1996.

XX 25-JAN-1996; 96WO-US01179.

XX 25-JAN-1995; 95US-0390301.

XX (CORT-) COR THERAPEUTICS INC.

XX Scarborough RM, Sundelin J;

XX WPI; 1996-362813/36.

XX N-PSDB; AAT32036.

XX

PT Vector for expression C140 cell surface receptor in host cell

PT useful to identify C140 agonist and antagonists, which are

PT antihypertensives and elevators of blood pressure, respectively

XX

PS Example 1; Fig 1A-B; 60pp; English.

XX

CC AAW01952 represents the murine C140 receptor (C140R), including a

CC putative signal peptide (see features table). DNA encoding C140R may be

CC engineered so as to allow the recombinant expression of C140R in a

CC suitable host cell, i.e. by removing the native expression-control

CC sequences and replacing them with control sequences operable in the

CC host. Such a recombinant receptor can be expressed on the surface of

CC oocytes, this provides a good assay system for identifying

CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked

CC receptor and a member of the "seven-pass" transmembrane receptor

CC superfamily (peptide chain of the receptor passes through the cell

CC membrane seven times, producing seven transmembrane regions within the

CC receptor molecule). The C140 receptor is involved in controlling blood

CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit

CC signalling from this receptor, resulting in an increase in blood pressure

CC and are therefore useful in pharmaceuticals for the treatment of

CC hypotension (low blood pressure). Conversely agonists (see

CC AAW01914-W01941) of C140 are useful in pharmaceuticals for the treatment

CC of hypertension (high blood pressure).

XX

SQ Sequence 395 AA;

Query Match 27.8%; Score 600.5; DB 17; Length 395;

Best Local Similarity 34.8%; Pred. No. 1.3e-53;

Matches 132; Conservative 69; Mismatches 115; Indels 63; Gaps 11;

Qy 30 LLLLLPTFCOSGMENDTNL--AKPTLPKTPRGAPPN---SPEEPFSALECGMTCATIT 84

Db 17 VMLLRLCTGRNNSKGRSLIGRLTQPTTGKGVPEPGFSIDEF----- 62

Qy 85 VKIKPEESASHLHVKNATMGYLTSLTKLIPAIYLLVGVGVPANAVTLMLFFRTRS 144

Db 63 -----SAS-----ILTGKLTTFVLPVWYIIVFVIGLPSNGMALWFLFRTKK 104

Qy 145 ICTTVFY-TNLAIADFLFCVTLFPKIAHYLNGNWNVFGVLCRATTVIFGNYMYSILL 203

Db 105 KHPAVIYMANLALADLLSVIFWFLKISYHLGNNWYVGEALCKVLIGFFYGNMYSILFM 164

Qy 204 ACISINRYLAIVHPFTYRGPL--KHTYALVTCGLVWATVFLYMLPFILKQEVYLVQPD 261

Db 165 TCLSVQRYWVIVNPM---GHPKPKANIAVGSLAIWLLFLVTIPDYVWKQIYIPALNI 221

Qy 262 ITTCHDVHNTCESSPPQLY-----YFISLAFPGFLIPFVLIYCYAAIIRTLNA----- 310

Db 222 ITTCHDV-----LPEEVLVGMFNYFLSLAIGVFLFPALLTASAVVLMIKTLRSSAMD 274

Qy 311 --YDHRWLWYVKASLLILVITTCFAPSNIILIIHHANYNNYNTDGLFYLIYALCLGSLN 369

Db 275 HSENKQRAIRLIITVLMYFICFRPSNLLLVVHYFLIKTQROSHVYALVLCISLTLN 334

Qy 370 NSCLDPFLYFLMSKT-RNHS 387

Db 335 NSCIDPFVYFVSKDFRDHA 353

CC same region of the long arm of chromosome number 5 (5q12-5q13)
CC as has been reported for the human thrombin receptor gene.
XX
SQ Sequence 398 AA;
Query Match 27.1%; Score 585.5; DB 16; Length 398;
Best Local Similarity 36.6%; Pred. No. 4.6e-52;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;
QY 80 GATITVKIKCPESASHLVKNATMGYLTSSLTSLKLIPIALYLLVFFVGVGPANAVTLWMLF 139
Db 53 GVTVEVFSVDEFSAS-----VLTKLTVPFPIVYTVFVGLPSNGMALWFL 102
QY 140 FTRISCTTVFY-TNLAIADFLCVTLPPFKIAYHLNGNNWVGEVLCRAITTVIFYGNMYC 198
Db 103 FRTKKKHPAIVYMANLADLLSVIWPPLKIAYHIGNNNWYGEALCNVLIGFFYGNMYC 162
QY 199 SILLACISINRYLAIVHPFTYRGLPKHYALVTCGLVWATVLYMLPFFILKQEVYLVQ 258
Db 163 SILFMTCLSVQRYWVIVNPMGH-SRKKANIAIGISLAIWLLILLVTPLYVVKQTIFIPA 221
QY 259 PDITTCCHDVNTCESSSPFOLY-----YFISLAFFGLPIPVLLIIVCYAAIIRTUNA-- 310
Db 222 LNIITTCADV-----LPEQLLVGDMFNYSLSAIGVFLPAPLTSAYVLMIRLSSA 274
QY 311 ----YDHRWLWVYKASLLILVITFCFAPSNIILIIHANYNNYNTDG---LYFIYLIAL 363
Db 275 MDENSEKKRKRAIKLIVTLVAMLYICTPSNLLLVH---YFLIKSGOGSHVYALIVAL 331
QY 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXXNDLRE 400
Db 332 CLSLNSCIDPFVYVYFVSHDFRDHAKNALLCRSVRTVKQ 370.
RESULT 9
AAW01953
ID AAW01953 standard; Protein; 398 AA.
XX
AC AAW01953;
DT
DT 01-APR-1997 (first entry)
XX
DE Human C140 receptor, with putative signal sequence.
XX
XX C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT /note= "putative signal peptide, differs from
FT this receptor (see AAW01955), the signal
FT sequence given for the cDNA clone is
FT believed to be the correct sequence"
FT
FT Protein 28..398
FT /note= "mature protein"
FT
FT Modified-site 31
FT /note= "potential Asn-linked glycosylation site"
FT
FT Cleavage-site 37..38
FT /note= "putative protease receptor cleavage site"
FT
FT Region 81..103
FT /note= "transmembrane region I"
FT
FT Region 111..132
FT /note= "transmembrane region II"
FT
FT Region 151..174
FT /note= "transmembrane region III"
FT
FT Region 191..212
FT /note= "transmembrane region IV"
FT
FT Modified-site 223
FT /note= "potential Asn-linked glycosylation site"
FT
FT Region 245..267

FT Region /note= "transmembrane region V"
FT 289..309
FT /note= "transmembrane region VI"
FT 327..348
FT /note= "transmembrane region VII"
XX
XX W096223225-A1.
PN 01-AUG-1996.
XX
XX 25-JAN-1996; 96W0-US01179.
XX
XX 25-JAN-1995; 95US-0390301.
XX (CORT-) COR THERAPEUTICS INC.
XX
XX Scarborough RM, Sundelin J;
XX
XX WPI; 1996-362813/36.
DR N-PSDB; AAT32037.
XX
PT Vector for expression C140 cell surface receptor in host cell
PT useful to identify C140 agonist and antagonists, which are
PT antihypertensives and elevators of blood pressure, respectively
XX
XX Example 2; Fig 2A-B; 60pp; English.
XX
XX AAW01953 represents the human C140 receptor (C140R), including a
XX putative signal peptide (see features table). DNA encoding C140R may be
XX engineered so as to allow the recombinant expression of C140R in a
XX suitable host cell, i.e. by removing the native expression-control
XX sequences and replacing them with control sequences operable in the
XX host. Such a recombinant receptor can be expressed on the surface of
XX oocytes, this provides a good assay system for identifying
XX agonists/antagonists of C140R. The C140 receptor is a G-protein linked
XX receptor and a member of the "seven-pass" transmembrane receptor
XX superfamily (peptide chain of the receptor passes through the cell
XX membrane seven times, producing seven transmembrane regions within the
XX receptor molecule). The C140 receptor is involved in controlling blood
XX pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
XX signalling from this receptor, resulting in an increase in blood pressure
XX and are therefore useful in pharmaceuticals for the treatment of
XX hypertension (low blood pressure). Conversely agonists (see
XX AAW01914-W01941) of C140 are useful in pharmaceuticals for the treatment
XX of hypertension (high blood pressure).
XX
XX Sequence 398 AA;
SQ
Query Match 27.1%; Score 585.5; DB 17; Length 398;
Best Local Similarity 36.6%; Pred. No. 4.6e-52;
Matches 124; Conservative 62; Mismatches 114; Indels 39; Gaps 10;
QY 80 GATITVKIKCPESASHLVKNATMGYLTSSLTSLKLIPIALYLLVFFVGVGPANAVTLWMLF 139
Db 53 GVTVEVFSVDEFSAS-----VLTKLTVPFPIVYTVFVGLPSNGMALWFL 102
QY 140 FTRISCTTVFY-TNLAIADFLCVTLPPFKIAYHLNGNNWVGEVLCRAITTVIFYGNMYC 198
Db 103 FRTKKKHPAIVYMANLADLLSVIWPPLKIAYHIGNNNWYGEALCNVLIGFFYGNMYC 162
QY 199 SILLACISINRYLAIVHPFTYRGLPKHYALVTCGLVWATVLYMLPFFILKQEVYLVQ 258
Db 163 SILFMTCLSVQRYWVIVNPMGH-SRKKANIAIGISLAIWLLILLVTPLYVVKQTIFIPA 221
QY 259 PDITTCCHDVNTCESSSPFOLY-----YFISLAFFGLPIPVLLIIVCYAAIIRTUNA-- 310
Db 222 LNIITTCADV-----LPEQLLVGDMFNYSLSAIGVFLPAPLTSAYVLMIRLSSA 274
QY 311 ----YDHRWLWVYKASLLILVITFCFAPSNIILIIHANYNNYNTDG---LYFIYLIAL 363
Db 275 MDENSEKKRKRAIKLIVTLVAMLYICTPSNLLLVH---YFLIKSGOGSHVYALIVAL 331
QY 364 CLGSLNSCLDPFLYFLMSKT-RNHS-TAYLTXXXNDLRE 400

QY 207 SINRYLAIHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQPDITTCND 266
Db 170 SVQRYWVIVNPMGH-SRKKANIAIGISLAIWLTLLVTLPIYVVKQTFIPALNITTCND 228
QY 267 VHTNCSSPPOLY-----YFISLAPFGFLIPFVLIYCYAAIIRTLNA-----YDHR 314
Db 229 V-----LPEQLLVGDMFNFLSLAIGVFLFAFLTASAYVLMIRMLRSSANDENSEKK 281
QY 315 WLWYKASLLILVIFTCFAPSNIILIHANYNNYNTDG---LYFIYLIALCLGSLNSC 371
Db 282 RKRAIKLIVTLVGLMYLICFTPSNLLVWH---YFLIKSQSQSHVYALYIVALCLSTLNSC 338
QY 372 LDPFLYFLMSKT-RNHS-TAYLTXXNDLREQQOP 404
Db 339 IDPFVYFVSHDFRDHAKNALLCRSVRTVKQMOP 373

RESULT 12

AAW01955
ID AAW01955 standard; Protein; 397 AA.

AC AAW01955;

XX 02-APR-1997 (first entry)

DT Human C140 receptor.

XX C140 receptor; G-protein linked; coupled; seven pass; agonist;

KW antagonist; hypertension; hypotension; blood pressure.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..27

FT /note= "the signal peptide differs from that

FT encoded by a genomic DNA sequence for

FT this receptor (see AAW01953), the signal

FT sequence given here is believed to be

FT the correct sequence"

FT Protein 28..397

FT /note= "mature protein"

XX WO9623225-A1.

XX 01-AUG-1996.

XX 25-JAN-1996; 96WO-US01179.

XX 25-JAN-1995; 95US-0390301.

XX (CORT-) COR THERAPEUTICS INC.

XX Scarborough RM, Sundelin J;

XX WPI; 1996-362813/36.

XX N-PSDB; AAT32039.

XX Vector for expression C140 cell surface receptor in host cell

PT useful to identify C140 agonist and antagonists, which are

PT antihypertensives and elevators of blood pressure, respectively

XX Example 5; Fig 11A-B; 60pp; English.

XX AAW01955 represents the human C140 receptor (C140R). DNA encoding C140R

CC may be engineered so as to allow the recombinant expression of C140R in

CC a suitable host cell, i.e. by removing the native expression-control

CC sequences and replacing them with control sequences operable in the

CC host. Such a recombinant receptor can be expressed on the surface of

CC oocytes, this provides a good assay system for identifying

CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked

CC receptor and a member of the "seven-pass" transmembrane receptor

CC superfamily (peptide chain of the receptor passes through the cell

CC membrane seven times, producing seven transmembrane regions within the

CC receptor molecule). The C140 receptor is involved in controlling blood
CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
CC signalling from this receptor, resulting in an increase in blood pressure
CC and are therefore useful in pharmaceuticals for the treatment of
CC hypotension (low blood pressure). Conversely agonists (see
CC AAW01914-W01941) of C140 are useful in pharmaceuticals for the treatment
CC of hypertension (high blood pressure).

XX Sequence 397 AA;

QY Query Match 26.4%; Score 569.5; DB 17; Length 397;

DB Best Local Similarity 34.2%; Pred. No. 2.1e-50;

QY Matches 135; Conservative 64; Mismatches 145; Indels 51; Gaps 13;

DB 28 AGLLLLPTFCQSGMENDTNNAKPTLPKTIKTRGAPPNSFEFPFSALEGWGTGATIVKI 87

DB 12 AAILLAASLSC-SGTIQGTNRSSKGRSLIGKVDGT-----SHVTG-KGVTVETVF 59

QY 88 KCPESASHLHVNATMGVLTSSLTSLKLIPIALYLLVGVGVANAVTLWMLPFRSICT 147

DB 60 SVDEFSSA-----VLAKLTTFVLPVITVIVFAVGLFSGNMGALWFLFRTKKHP 109

QY 148 TVFY-TNLAIADFLFCVTLPEKIAYHLGNVNVFGEVLCRAITVIFYGMYCSILLACI 206

DB 110 AVIYMANLALADLLSVIFPLKIAIHGNNIYGALCNVLIGFFRMYCSILFTTCL 169

QY 207 SINRYLAIHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQPDITTCND 266

DB 170 SVQRYWVIVNPMGH-SRKKANIAIGISLAIWLTLLVTLPIYVVKQTFIPALNITTCND 228

QY 267 VHTNCSSPPOLY-----YFISLAPFGFLIPFVLIYCYAAIIRTLNA-----YDHR 314

DB 229 V-----LPEQLLVGDMFNFLSLAIGVFLFAFLTASAYVLMIRMLRSSANDENSEKK 281

QY 315 WLWYKASLLILVIFTCFAPSNIILIHANYNNYNTDG---LYFIYLIALCLGSLNSC 371

DB 282 RKRAIKLIVTLVGLMYLICFTPSNLLVWH---YFLIKSQSQSHVYALYIVALCLSTLNSC 338

QY 372 LDPFLYFLMSKT-RNHS-TAYLTXXNDLREQQOP 404

DB 339 IDPFVYFVSHDFRDHAKNALLCRSVRTVKQMOP 373

RESULT 13

AAV15081

ID AAV15081 standard; Protein; 396 AA.

XX AAV15081;

XX 25-JAN-2000 (first entry)

XX Mouse protease-activated receptor 4 (PAR4).

XX Mouse protease-activated receptor 4; PAR4; thrombin receptor;

XX phosphoinositide hydrolysis; calcium efflux; platelet aggregation;

XX platelet activation; thrombin agonist; thrombin antagonist; therapeutic;

XX wound; blood coagulation; heart attack; stroke; inflammatory response;

XX proliferative response; atherosclerosis; restenosis; glomerulosclerosis;

XX pulmonary inflammation; ARDS; adult respiratory distress syndrome.

XX Mus musculus.

XX Key Location/Qualifiers

XX Cleavage-site 59..60

XX /label= Thrombin_cleavage_site

XX Peptide 60..65

XX /label= PAR4_activating_peptide

XX Protein 1..396

XX /label= PAR4_protein

XX /note= "The sequence is stated to be 397 amino acids

XX long in the specification but is shown to have 396 aa

XX in figure-2 as well as Sequence Listing"

XX


```
CC signal/FLAG epitope peptide shown on page 64.
SQ Sequence 408 AA;

Query Match 24.3%; Score 525; DB 20; Length 408;
Best Local Similarity 33.6%; Pred. No. 8.9e-46;
Matches 123; Conservative 66; Mismatches 157; Indels 20; Gaps 7;

QY 23 LIFAAAGLLLLLPTFCQ---SGMENDTNNAKPTLPIKTF--RGAPPNSPEEFP--FSAL 75
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15 LLLVWSNLLL-----CQGVSDYKDDDDKLEGGTQTPSVYDESGTGGDDSTPSILPAP 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 76 EGWTGATITVKIKCPESASHLVKNATMGYLTSLSLTKLIPAIYLLVFWGVPAANAVTL 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 70 RGYPGQV-----CANDSDT--LELPDSSRALLLGMVETRLVPALYGLVVLVPLANGAL 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 136 WMLFFRTSRISCTTVFYNNLAIADFLCVTLFPFKIAVHLNGNNVFGVLCRATTIVFYGN 195
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 WVLATQAPRLPSTWMLNLTATDLLLALALPPIAIVHLRQRPFGGAACRLATAALYGH 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 196 MYCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYY 255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 183 MYGSVLLAAVSLDYLALVPLRARALRGRLALGCMAAWMAAALALPLTLRQTFR 242
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 256 LVOPDITTCDDVNTCESSPFQLYFISLAFTGLIPFVLLIYCVAIIRTLNAYDHRW 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 243 LARSDRVLCALDPLDQAQSHQ--PAFTCLALLGCFLLPMLLCYATLHTLAASGRRY 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 316 LWTVKASLLLVITFCFAPSNIILIIHANYNNYNTDGLYFIYLIATCLGSLNSCLDPF 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 GHALRLTAVVLASAVFAVFSNLLLLLHYSDPSPSAWGNLYGAYVPSLALSTLNSCVDPF 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 376 LYFILMS 381
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 IYYVVS 367
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
AAV50139
ID AAV50139 standard; Protein; 385 AA.
XX
AC AAV50139;
XX
DT 31-JAN-2000 (first entry)
XX
DE Mutant human protease-activated receptor PAR4, R68A.
XX
KW Protease-activated receptor; PAR4; G protein coupled;
KW cellular signalling; protease; tethered ligand; N-terminal;
KW proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist;
KW antagonist; cellular response; physiological response; clotting pathway;
KW platelet; proliferation; differentiation; mediation;
KW inflammatory process; vascular injury; chemotaxis; mitogenesis;
KW growth factor; production; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..78 "Extracellular N-terminal region"
FT Peptide 1..17
FT "Signal peptide"
FT Cleavage-site 17..18
FT /note= "Cleared by signal peptidase"
FT Protein 18..385
FT /note= "Mature non-activated human PAR4"
FT Region 48..53
FT /note= "Tethered hexapeptide activating ligand"
FT (AAV50140) "Modified-site 56"
FT /note= "N-glycosylated"
FT Domain 79..102
FT /note= "Transmembrane domain 1"
```

FT Region

103..110

/note= "Intracellular loop 1"

FT Domain

111..132

FT Region

/note= "Transmembrane domain 2"

FT Region

133..150

FT Domain

/note= "Extracellular loop 1"

FT Region

151..172

FT Domain

/note= "Transmembrane domain 3"

FT Region

173..191

FT Domain

/note= "Intracellular loop 2"

FT Region

192..213

FT Domain

/note= "Transmembrane domain 4"

FT Region

214..239

FT FT

/note= "Extracellular loop 2, determines specificity for activating peptide"

FT Region

228..230

FT Domain

/note= "These three residues are conserved among PAR1-4"

FT Domain

240..263

FT Region

/note= "Transmembrane domain 5"

FT Region

264..283

FT Domain

/note= "Intracellular loop 3"

FT Domain

284..305

FT Region

/note= "Transmembrane domain 6"

FT Region

306..316

FT Domain

/note= "Extracellular loop 3"

FT Region

317..343

FT Domain

/note= "Transmembrane domain 7"

FT Region

344..385

FT FT

/note= "Intracellular C-terminal region"

XX

PN

WO9950415-A2.

XX

PD

07-OCT-1999.

XX

PF

31-MAR-1999; 99WO-US07100.

XX

PR

01-APR-1998; 98US-0053866.

XX

PA

(ZYMO) ZYMOGENETICS INC.

PA

(UNIW) UNIV WASHINGTON.

XX

PI

Xu W, Presnell SR, Yee DP, Foster DC;

XX

DR

WPI; 1999-633640/54.

XX

XX

PT

Novel protease activated receptor 4, useful for screening for

PT

(antagonists for promoting the proliferation and/or differentiation of

PT

platelets and in mediating inflammatory events -

XX

PS

Example 2; Page -; 85pp; English.

XX

CC

This sequence represents a mutant human protease-activated receptor PAR4

CC

R68A, compared with PAR4 mutant R47A (AAV50138) in studies to evaluate

CC

the importance of a protease cleavage site between residues 47 and 48.

CC

This mutant PAR4 and the wild-type were able to be cleaved by

CC

thrombin or trypsin, in contrast to the R68A mutant. This cleavage

CC

activates PAR4, exposing a tethered hexapeptide ligand at the new

CC

N-terminus. Protease-activated receptors (PARs) are a subfamily of G

CC

protein coupled receptors which are capable of mediating cellular

CC

signalling in response to proteases. Agonists of PAR4 are useful for

CC

upregulating cellular or physiological responses whereas antagonists

CC

are used to downregulate these activities. The PAR4 protein is

CC

further useful for dissecting the effects of thrombin or other

CC

activating proteases in the clotting pathway from the effects of these

CC

proteases at the cellular level. Agonists are specifically useful in

CC

promoting the proliferation and/or differentiation of platelets, in

CC

mediating inflammatory events, responses to vascular injury, chemotaxis

CC

or mitogenesis, and in producing growth factors. Antagonists are useful

CC

as research reagents for characterising sites of ligand-receptor

CC

interaction.

CC

Note: This sequence is not shown in the specification, but is derived

CC

from the sequence of the PAR4 shown in figure 1.

XX

SQ Sequence 385 AA;
 Query Match 24.1%; Score 523.5; DB 20; Length 385;
 Best Local Similarity 33.4%; Pred. No. 1.2e-45;
 Matches 123; Conservative 62; Mismatches 142; Indels 41; Gaps 6;
 QY 29 GLILLLP-----TFQSGMENDTNLAKPTLPKTPRGAPPNSFEFFPFS 73
 DB 3 GRLLLPVLGSLGGTQTPSYVDESGTGDDSTPSILPAP--RGYP-----50
 QY 74 ALGWTGATTVIKKPEBSASHLVKNATMGVLTSSLTKLIPAIYLLVVFVGVANAV 133
 DB 51 ---GQVCANDSTLELPDSSAALL-----LGV-----VPTRLVPALYGLVVLVGLPANG 97
 QY 134 TLAWLPFRTRISCTTVFYTNLATADFLCVTLFPKTIAYHLNGNNVVFGEVLGRATTVIF 193
 DB 98 ALWVLTQAPRLPSTMLMNLATADLLALPRLIAYHLRGQRPFGAACLATAALY 157
 QY 194 GNNYCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKOE 253
 DB 158 GHMYGSVLLLAAYSLDRYLALVHLPLRARALGRRLALGLCMAAWLMAALALPLTLQRT 217
 QY 254 YVLVQPDITTCVHDVNTCESSPFQLYYFISLAFGLFIPFVLIYCYAAIIRTNAYDH 313
 DB 218 FRLARSDRVLCHDALPLDAQSHWQ-PAFTCLALLGCFPLPLAMLLCYGATLHTLAASGR 276
 QY 314 RLLWYVKASLLIIVITFCAPSNIILIIHHANYNNYNTDGLYFIYLIACLGSLNSCLD 373
 DB 277 RYGHAKRLTAVLASAVFVPSNLLLLHYDPSFSAWGNLYGAYVPSLALSTLNSCVD 336
 QY 374 PFLYFLMS 381
 DB 337 PFIYIYVS 344

RESULT 16

AAR27240
 ID AAR27240 standard; Protein; 425 AA.

XX AAR27240;
 AC AAR27240;

XX 21-MAY-1998 (first entry)

XX Human thrombin receptor (TR).

XX Diagnosis; cardiovascular disease; wound healing; restenosis;
 KW thrombosis; unstable angina treatment; myocardial infarction;
 KW thrombotic; thromboembolytic stroke.

XX Homo sapiens.

Key Location/Qualifiers
 FT peptide 1..26
 FT /label= sig_peptide
 FT peptide 28..425
 FT /label= mat_protein
 FT peptide 28..41
 FT /note= "activation peptide RPEKATNATLDR"
 FT cleavage_site 41..42
 FT /notes "thrombin-catalysed cleavage-site"
 FT disulfide_bond 175..254
 FT /note= "analogous to rhodopsin and beta-2
 FT adrenergic receptor"
 FT domain 100..129
 FT /notes "transmembrane domain I"
 FT domain 137..160
 FT domain 177..197
 FT /note= "transmembrane domain II"
 FT domain 219..239
 FT /note= "transmembrane domain III"
 FT domain 266..294
 FT /note= "transmembrane domain IV"
 FT domain 304..344
 FT /note= "transmembrane domain V"

FT domain 311..335
 FT /note= "transmembrane domain VI"
 FT domain 348..367
 FT /note= "transmembrane domain VII"
 FT region 34..36
 FT /note= "consensus N-linked glycosylation site"
 FT region 62..64
 FT /note= "consensus N-linked glycosylation site"
 FT region 75..77
 FT /note= "consensus N-linked glycosylation site"
 FT region 250..252
 FT /note= "consensus N-linked glycosylation site"
 FT region 259..261
 FT /note= "consensus N-linked glycosylation site"
 XX WO9214750-A.
 PN 03-SEP-1992.
 XX 19-FEB-1992; 92WO-US01312.
 XX 19-FEB-1991; 91US-0657769.
 PR 07-NOV-1991; 91US-0789184.
 XX (CORT-) COR THERAPEUTICS INC.
 PA (REGC) UNIV CALIFORNIA.
 XX Coughlin SR, Scarborough RM;
 PI WPI; 1992-316119/38.
 DR N-PSDB; AAQ28568.
 XX DNA encoding cell surface receptor for thrombin - useful for
 PT determining thrombin in diagnosing e.g. cardiovascular diseases,
 PT also to treat wound healing, restenosis etc.
 XX Disclosure; Fig 1; 81pp; English.
 XX The TR DNA sequence can be used in the prepn. of diagnostics to
 CC determine thrombin levels in samples, and screening tools for
 CC candidate substances which affect thrombin activity in vivo.
 CC Thrombosis may be diagnosed in a mammal by measuring the presence,
 CC absence or amt. of the cleaved activation peptide of the TR.
 XX Sequence 425 AA;
 Query Match 24.2%; Score 523; DB 13; Length 425;
 Best Local Similarity 32.9%; Pred. No. 1.5e-45;
 Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;
 QY 14 GTQVIKMKALIFAAAGLLLLPTFCQSGMENDTNLAKPTLPKTPRGAPPN-SREFFPF 72
 DB 2 GPRLLLVAAACFSLCGLPLLSARTRRPPESKATN----ATLDRPSFLLRNPNDKYEPFWE 57
 QY 73 SALEGWGTGAT---ITVKIKPEBSASHLVKNATMGVLTSSLTKLIPAIYLLVVFVGV 128
 DB 58 DEEKNEGLTEYRLVSINKSSPLQKLPAPFISDASGYLTSSWLTLPVPSVVTGVFVSL 117
 QY 129 PAN--AVTLWMLFFRTRISCTTVFYTNLAIAFLPCVTLPPKTIAYHLNGNNVVFGEVL 186
 DB 118 PLNIMAIVVFIKMKVKK-PAVVYMLHLATADLVFVSLPFPKISYVFGSDMQFSEL 176
 QY 187 ATTVIFYGNMYSILLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVWATVFLY 243
 DB 177 FVTAAPFCNMYSILLMTVISIDRLAVVYPMQUSWRTLGRASP---TCLAIWALAIA 233
 QY 244 MLFPFLKQEVYLVQPDITTCVHDVNTCESSPFQLYYFISLAFGLFIPFVLIYCYAA 303
 DB 234 VVPLVLKEQTIQVPGNLITTCVHDVNTLLEG-VYAYFSAFSAFVFFPLIISTVCYVS 292
 QY 304 IIRTLN-----AYDHRWLWYVKASLLILVITTCFAPSNIILIIHHANY-YNNVTDGL 355
 DB 293 IIRCLSSAVANRKKRALF---LSAAVFCIFICFGPTNVLLIAHYSFLSHTSTTEA 349


```

QY      356 YFIYLIALCGSLNSCLDPEFLYFLMS 381
DB      177 FVTAAPFCNMYASILMTVISIDRFVAVYPMQSLSWRTLGRASF---TCLAIWALAIAAG 233

RESULT 17
AAW51407
ID      AAW51407 standard; Protein; 425 AA.
XX      AC
XX      AAW51407;
DT      12-OCT-1998 (first entry)
DE      Human protease-activated receptor 1 (PAR1).
XX      KW
XX      KW
XX      KW
XX      OS Homo sapiens.
XX      FH
XX      FH
XX      FT Binding-site 52..55
XX      FT /note= "thrombin binding site"
XX      FT Cleavage-site 41..42
XX      FT /note= "thrombin cleavage site"
XX      PN
XX      PN WO9818456-A1.
XX      PD
XX      PD 07-MAY-1998.
XX      PF
XX      PF 29-OCT-1997; 97WO-US19732.
XX      PR
XX      PR 30-OCT-1996; 96US-0742440.
XX      PA (REGC ) UNIV CALIFORNIA.
XX      PI
XX      PI Connolly A, Coughlin SR, Ishihara H;
XX      DR WPI; 1998-271905/24.
XX      PT
XX      PT DNA encoding protease-activated receptor 3 - for detection of
XX      PT specific agonists and antagonists, potentially useful for treating
XX      PT e.g. thrombosis, atherosclerosis, inflammation etc.
XX      PS
XX      PS Example 1; Page 42-43; 74pp; English.
XX      CC
XX      CC This polypeptide comprises human protease-activated receptor 1
XX      CC (PAR1), a receptor that mediates thrombin signalling. The
XX      CC invention relates to novel mouse and human PAR3 (see AAW51405-06)
XX      CC that show homology to PAR1 and which are specific receptors for
XX      CC thrombin. They can be used to screen for specific agonists and
XX      CC antagonists of thrombin useful e.g. for treating atherosclerosis,
XX      CC thrombosis and inflammation.
XX      SQ
XX      SQ Sequence 425 AA;
XX      Query Match 24.2%; Score 523; DB 19; Length 425;
XX      Best Local Similarity 32.9%; Pred. No. 1.5e-45;
XX      Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;
QY      14 GTQVIMKALIFAAAGLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPN-SFEFFPF 72
DB      2 GPRLLLVACFSLCGLLSARTARPESKATN----ATLDRSFLLRNPNDKYEPFWE 57

QY      73 SALEGWGTAT-----ITVKIKCPESASHLVKQATMGYLTSSISTKLIPAIYLLVFVGV 128
DB      58 DEEKNESGLTEYRLVLSINKSPLOKQLPAFISEDASGYLTSSWLTLPVPSVITGVFVSL 117

QY      129 PAN--AVTLWMFLPRTSICITTFYTNLAITADELFCVTLPFKIAYHLNGNNWVGEVLOR 186
DB      118 PLNIMAIVFPMKMKVKK--PAVVMYMLHLATADLVFVSLVPFKISYFSGSDWQFSELOR 176

QY      187 ATTVPFYGNMYCSILLACISINRYLAIVHP---FTYRGLPKHYALTVCGLVWATVFLY 243
DB      177 FVTAAPFCNMYASILMTVISIDRFVAVYPMQSLSWRTLGRASF---TCLAIWALAIAAG 233

QY      244 MLPFFILKQEYLVQPDITTCDDVHNTCESSSPFOLYFISLAFFGLIPFVLLIYCYAA 303
DB      234 VVPLVLEQTIQVGLNITTCDDVHNTLLEG-YVAYFSAFSAVFFVPLIISTVCYVS 292

QY      304 IIRTLN-----AYDHRWLWYKASLLILVIFTTCFAPSNIILIIHHANY-YNNNTDGL 355
DB      293 IIRCLSSAVANRSKSRALF---LSAAVFCIFIIICGPTNVLLIAHYSLTSHTSTEEA 349

QY      356 YFIYLIALCGSLNSCLDPEFLYFLMS 381
DB      350 YFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 18
AAW49570
ID      AAY49570 standard; Protein; 425 AA.
XX      AC
XX      AAY49570;
DT      13-JAN-2000 (first entry)
XX      DE
XX      DE Human thrombin receptor protein sequence.
XX      KW
XX      KW Human; coding sequence polymorphism; vascular pathology gene;
XX      KW polymorphic site; phenotype correlation; forensic; paternity testing;
XX      KW medicine; genetic analysis; vascular disease.
XX      OS
XX      OS Homo sapiens.
XX      PN
XX      PN WO9950454-A2.
XX      PD
XX      PD 07-OCT-1999.
XX      PF
XX      PF 26-MAR-1999; 99WO-US06473.
XX      PR
XX      PR 01-APR-1998; 98US-0054272.
XX      PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX      PI
XX      PI Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;
XX      DR WPI; 1999-620066/53.
XX      DR N-PSDB; AA232191.
XX      PT
XX      PT Determination of polymorphisms in genes, especially those identifying
XX      PT predisposition to vascular disease
XX      PS
XX      PS Disclosure; Fig 35; 134pp; English.
XX      CC
XX      CC AA232159 to AA232194 represent reference alleles for specifically
XX      CC claimed nucleic acid sequences from the present invention which comprise
XX      CC polymorphic sites as given in a table in the specification, selected
XX      CC from 92 single nucleotide polymorphisms in which the nucleotide at the
XX      CC polymorphic site is different from a nucleotide at the same site in a
XX      CC reference allele. The nucleic acids, and primers and probes, are used to
XX      CC identify polymorphisms, which may predispose an individual to disease,
XX      CC especially a vascular disease. They can also be used in phenotype
XX      CC correlations, forensics, paternity testing, medicine or genetic
XX      CC analysis. AAY49550 to AAY49573 represent the proteins which correspond
XX      CC to some of the reference alleles.
XX      SQ
XX      SQ Sequence 425 AA;
XX      Query Match 24.2%; Score 523; DB 20; Length 425;
XX      Best Local Similarity 32.9%; Pred. No. 1.5e-45;
XX      Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;
QY      14 GTQVIMKALIFAAAGLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPN-SFEFFPF 72
DB      118 PLNIMAIVFPMKMKVKK--PAVVMYMLHLATADLVFVSLVPFKISYFSGSDWQFSELOR 176

```

Db 2 GPRRLLLVAACFSLCGLPILLSARTRARRPESKATN-----ATLDRPSFLLRNPNDKYPFW 57
QY 73 SALEGWGAT----ITVKIKCPESASHLHVKNATMGYLTSSSTKLIPAIYLLVFGV 128
Db 58 DEEKESGLTEYRLVSNKSPLOKLPAFISEDAGYLTSSWLTUFPVSVYTGVPVWSL 117
QY 129 PAN--AVTLWMLFFRTRISCTTVFYTNLAIADEFLFCVTLPFKIAYHLNGNNWVFGVLCR 186
Db 118 PLNIMAVVFLKMKVKK-PAVYMLHLATADLVFVSVLPFKISYFSGSDMQFGSEL 176
QY 187 ATTVPYGNMYCISILLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLWATVFLY 243
Db 177 FVTAAPFCNNYASILLMTVISIDRFVAVYPMQSLSWRTLGRASF---TCLAIWALAIA 233
QY 244 MLPPFLLKQEVYLVQPDITTCVDVHNTCESSPPQLYYFISLAFPGFLIPFVLIICYAA 303
Db 234 VVPLVLKEQTIQVGLNITTCVDVHNTLLEG-YYAYFSAFSAVFFVPLIISTVCYVS 292
QY 304 IIRTLN-----AYDHRWLWYVKASLLILVITTCFAPSNIILIIHHANY-YNNNTDGL 355
Db 293 IIRCLSSAVANRKSRAFP---LSAAVCFIICFGPTNVLLIAHYSFLSHTSTEEA 349
QY 356 YFIYLIACLGSLNSCLDPFLYFLMS 381
Db 350 YFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 19

AAG80697
ID AAG80697 standard; Protein; 425 AA.

XX AC AAG80697;

XX DT 14-MAR-2002 (first entry)

XX DE Human thrombin protein.

XX KW G-protein coupled receptor; human; analgesic; antirheumatic; norepinephrine;
KW antidiarrhetic; neuroprotective; antiparkinsonian; hypotensive; cardiac;
KW antidiarrhetic; antigravine; cyostatic; antidiabetic; cardiac disorder;
KW muscular activity; immunosuppressive; vasotropic; gene therapy; cancer;
KW cellular processes modulation; intracellular signalling modulation;
KW proliferative disorder; central nervous system disorder; CNS; pain;
KW metabolic disorder; muscular disorder; rheumatoid arthritis; depression;
KW Alzheimer's disease; Parkinson's disease; autonomic function disorder;
KW hypertension; neuropsychiatric disorder; learning disorder;
KW memory disorder; bipolar effective neurological disorder;
KW diabetes mellitus; autoimmune disorder; thrombin.

XX OS Homo sapiens.

Key	Location/Qualifiers
Domain	102..132
FT	/label= TW1
FT	/note= "Transmembrane domain 1"
Domain	138..165
FT	/label= TW2
FT	/note= "Transmembrane domain 2"
Domain	176..200
FT	/label= TW3
FT	/note= "Transmembrane domain 3"
Domain	215..241
FT	/label= TW4
FT	/note= "Transmembrane domain 4"
Domain	268..296
FT	/label= TW5
FT	/note= "Transmembrane domain 5"
Domain	313..340
FT	/label= TW6
FT	/note= "Transmembrane domain 6"
Domain	347..379
FT	/label= TW7
FT	/note= "Transmembrane domain 7"

XX PN WO200196400-A2.
XX PD 20-DEC-2001.
XX PF 18-JUN-2001; 2001WO-US19464.
XX PR 16-JUN-2000; 2000US-212331P.
XX PR 16-FEB-2001; 2001US-269758P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Glucksmann MA, Silos-Santiago I;
XX WPI; 2002-098057/13.
XX PT New isolated human G protein-coupled receptor polypeptide, 52871, for
PT diagnosing and treating proliferative, central nervous system, pain,
PT cardiac, metabolic or muscular disorders -
XX PS Disclosure; Page 106-107; 112pp; English.
XX CC This invention describes a novel isolated G protein-coupled receptor
CC (GPCR) polypeptide (I), designated 52871. The product of the invention
CC has analgesic, antirheumatic, antidiarrhetic, norepinephrine, neuroprotective,
CC antiparkinsonian, hypotensive, antidiabetic, antigravine, cyostatic,
CC antidiabetic, immunosuppressive, vasotropic, cardiac, muscular activity
CC and can be used for gene therapy, cellular processes modulation and
CC intracellular signalling modulation. (I) is useful in screening assays, predictive
CC or an antibody derived from (I) is useful in screening assays, predictive
CC medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical
CC trials and pharmacogenomics), and in methods of treatment (e.g.
CC therapeutic and prophylactic). (I) or (II) are useful for treating
CC 52871-associated disorders such as proliferative disorders, central
CC nervous system disorders, pain or pain disorders, cardiac disorders,
CC metabolic disorders or muscular disorders. (I) or (II) are useful for
CC treating pain disorders such as rheumatoid arthritis, CNS disorder such
CC as Alzheimer's disease, Parkinson's disease, autonomic function disorders
CC such as hypertension, neuropsychiatric disorders such as depression,
CC learning or memory disorders, bipolar effective neurological disorders
CC such as migraine, cardiac-related disorders such as stenosis, hormonal
CC disorders such as diabetes mellitus, autoimmune disorders and cancer.
CC This sequence represents the human thrombin protein described in the
CC method of the invention.

XX SQ Sequence 425 AA;

Query Match 24.2%; Score 523; DB 23; Length 425;
Best Local Similarity 32.9%; Pred. No. 1.5e-45;
Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;

QY	14	GTQVIMKALIFAAAGLLLLLPFCQSGMENDTNNAKPTLPKIPRAGPN-SPEEPFF	72
Db	2	GPRRLLLVAACFSLCGLPILLSARTRARRPESKATN-----ATLDRPSFLLRNPNDKYPFW	57
QY	73	SALEGWGTAT----ITVKIKCPESASHLHVKNATMGYLTSSSTKLIPAIYLLVFGV	128
Db	58	DEEKESGLTEYRLVSNKSPLOKLPAFISEDAGYLTSSWLTUFPVSVYTGVPVWSL	117
QY	129	PAN--AVTLWMLFFRTRISCTTVFYTNLAIADEFLFCVTLPFKIAYHLNGNNWVFGVLCR	186
Db	118	PLNIMAVVFLKMKVKK-PAVYMLHLATADLVFVSVLPFKISYFSGSDMQFGSEL	176
QY	187	ATTVPYGNMYCISILLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLWATVFLY	243
Db	177	FVTAAPFCNNYASILLMTVISIDRFVAVYPMQSLSWRTLGRASF---TCLAIWALAIA	233
QY	244	MLPPFLLKQEVYLVQPDITTCVDVHNTCESSPPQLYYFISLAFPGFLIPFVLIICYAA	303
Db	234	VVPLVLKEQTIQVGLNITTCVDVHNTLLEG-YYAYFSAFSAVFFVPLIISTVCYVS	292
QY	304	IIRTLN-----AYDHRWLWYVKASLLILVITTCFAPSNIILIIHHANY-YNNNTDGL	355

Db 293 IIRCLSSAVANRKKSRALF---LSAAVFCFIICFGPTNVLLIAHYFLSHTSTTEAA 349

QY 356 YFIYLIACLGSLNSCLDPFLYFLMS 381

Db 350 YFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 20

AAE17032

ID AAE17032 standard; Protein; 425 AA.

AC AAE17032;

XX 18-APR-2002 (first entry)

DT Human thrombin protein.

DE Human; G protein-coupled receptor; GPCR; SLGP; cellular proliferation;

XX growth; differentiation; migration disorder; cancer; retinal; vasotrophic;

KW optic disk neovascularisation; cytostatic; gene therapy; genetic disease;

KW tissue typing; forensic biology; antiarthritic; ischaemia; arthritis;

KW thrombin.

XX Homo sapiens.

OS

XX

Key Location/Qualifiers

Domain 102..132

FT /label= TM1

FT /note= "Transmembrane domain"

Domain 138..165

FT /label= TM2

FT /note= "Transmembrane domain"

Domain 176..200

FT /label= TM3

FT /note= "Transmembrane domain"

Domain 215..241

FT /label= TM4

FT /note= "Transmembrane domain"

Domain 268..296

FT /label= TM5

FT /note= "Transmembrane domain"

Domain 313..340

FT /label= TM6

FT /note= "Transmembrane domain"

Domain 347..379

FT /label= TM7

FT /note= "Transmembrane domain"

XX

PN WO200202602-A2.

XX

10-JAN-2002.

XX

29-JUN-2001; 2001WO-US20751.

XX

30-JUN-2000; 2000US-0608921.

XX

(MILL-) MILLENNIUM PHARM INC.

XX

TSai F;

XX

WPI; 2002-140083/18.

XX

New isolated G protein-coupled receptor nucleic acid molecule, SLGP,

XX useful for diagnosing and treating cellular proliferation, growth,

PT differentiation, or migration disorders e.g. cancer, arthritis and

PT myocardial ischemia

XX

Disclosure; Page 158-159; 182pp; English.

XX

The present invention relates to an isolated G protein-coupled receptor

CC (GPCR) nucleic acid molecule, SLGP which encodes a polypeptide. SLGP is

CC useful as target or therapeutic agent for diagnosis and treatment of

CC cellular proliferation, growth, differentiation, or migration disorders

CC (e.g., cancer, arthritis, retinal and optic disk neovascularisation, and

CC tissue ischaemia, such as myocardial ischaemia). SLGP is useful in

CC screening assays, detection assays (e.g., chromosomal mapping, tissue

CC typing, forensic biology), predictive medicine (e.g., diagnostic assays,

CC prognostic assays, monitoring clinical trials and pharmacogenomics), and

CC in methods of treatment (e.g., therapeutic and prophylactic). SLGP is

CC useful as reagents or targets in assays applicable to treatment and

CC diagnosis of SLGP-mediated or related disorders. SLGP is useful in gene

CC therapy, to express SLGP protein, to detect SLGP mRNA or a genetic

CC alteration in a SLGP gene, and to modulate SLGP activity. SLGP is useful

CC to map their respective genes on a chromosome, and thus locate gene

CC regions associated with genetic disease, to identify an individual from a

CC minute biological sample (tissue typing), and to aid in forensic

CC identification of a biological sample. The present sequence is human

XX thrombin protein used in the invention.

XX

SQ Sequence 425 AA;

Query Match 24.2%; Score 523; DB 23; Length 425;

Best Local Similarity 32.9%; Pred. No. 1.5e-45;

Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;

QY 14 GTQVIKMKALIPAAAGLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPN-SFEFFPF 72

Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 73 SALEGWTGAT---ITVKKICPEESASHLHVKNATMGYLTSSLTSLKLIPIAYLLLVFVGV 128

Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 58 DEEKNEGLTEYRLVLSINKSSPQKQLPAPFISDASGYLTSSWLTLPVPSVYTGTVFVSL 117

Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 129 PAN--AVTLWMLFFRTRISCTTVFYTNLAIAIDFLFCVTLFPFKIAYHLNGNNWVGEVL 186

Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 118 PLNIMAVVFLKMKVKK-PAVVYMLHLATADVFLSVLPFKISYIFSGSDWQFSEL 176

QY 187 ATTIVFYGNMYCSILLACISINRYLIVHP---FTYRGLPKHYALVTCGLVNAIVFLY 243

Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 177 FVTAAPFCNMYSILLMTVISIDRFLAVVYPMQSLWSRTLGRASF--TCLAIWALAIAG 233

QY 244 MLPFFILKQYVLVQPDITTDHVNHTCESSPFQLYYFISLAFFGLPFPVLIIVCYAA 303

Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 234 VVPLVKEQTIQVPGNITTDHVNHTLLEG-YAYYFSAFSAVFFVPLIISTVCYVS 292

QY 304 IIRTLN-----AYDHRWLWYVKASLLILVITFCAPSNIILIIHHANY-YNNNTDGL 355

Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 293 IIRCLSSAVANRKKSRALF---LSAAVFCFIICFGPTNVLLIAHYFLSHTSTTEAA 349

QY 356 YFIYLIACLGSLNSCLDPFLYFLMS 381

Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

QY 350 YFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 21

AAE17032

ID AAY45035 standard; Protein; 426 AA.

XX

AC AAY45035;

XX

31-MAY-2000 (first entry)

XX

Human thrombin receptor, a member of protease activated receptor family.

XX

Human; thrombin receptor; THR; PAR-1; protease activated receptor;

KW antiseptic molecule; PAR antibody; cytostatic; therapeutic;

KW metastatic tumour cell; placental implantation; invasive cell.

XX

OS Homo sapiens.

XX

Key Location/Qualifiers

FT 37..61

FT /note= "conserved unique region comprising protease

FT activated domain and hirudin binding domain.

FT Therapeutically useful antiseptic molecules are designed

FT based on this region"

FT Peptide 42..51
 FT /label= TRAP
 FT /note= "Thrombin-receptor activating peptide used
 FT for producing anti-Thr antibodies"
 FT Misc-difference 426
 FT /note= "Encoded by TAG"
 FT Region 251..257
 FT /note= "conserved residues in second extracellular
 FT loop used for designing therapeutically useful
 FT antisense molecules "
 XX
 PN WO200008150-A1.
 XX
 PD 17-FEB-2000.
 XX
 XX 05-FEB-1999; 99WO-1100079.
 PF
 XX 07-AUG-1998; 98IL-0125698.
 PR
 XX (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
 PA
 XX Bar-Shavit R;
 PI
 XX WPI; 2000-205706/18.
 DR N-PSDE; AAZ50771.
 DR
 XX Treating metastatic tumor cells useful for treating disorders involving
 XX placenta implantation in a female comprises administration of an
 XX antisense molecule complementary to an RNA sequence of a protease
 XX activated receptor protein -
 XX
 PS Example 2; Fig 1b; 46pp; English.
 XX
 CC The patent discloses a method to treat metastatic tumour cells using
 CC an antisense molecule comprising a polynucleotide complementary to an
 CC RNA sequence of a protease activated receptor (PAR) protein, or an
 CC antibody capable of binding to a PAR protein. The antisense molecules and
 CC antibodies of PAR protein are also used to treat disorders associated
 CC with implantation of placenta. The present sequence is a
 CC human thrombin receptor (THR) (also known as PAR-1), which is a
 CC G-coupled protein belonging to the PAR family. This sequence is used to
 CC produce antisense molecules or antibodies against PAR protein for
 CC therapeutic use.
 XX
 SQ Sequence 426 AA;
 Query Match 24.2%; Score 523; DB 21; Length 426;
 Best Local Similarity 32.9%; Pred. No. 1.5e-45;
 Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;
 QY 14 GTQVIRKALIFAAGLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPN-SFEEPPF 72
 DB 2 GPRLLVAACFLCGPLLSARTARRPESKATN---ATLDRPFLLRNPDKPFFWE 57
 QY 73 SALEGWTGAT----ITVKIKCPESASHLVKNATMGYLTSSLSLTKLPAIYLLFVGVG 128
 DB 58 DEEKNESGLTEYRLVLSINKSPQKQLPAFISEDASGYLTSSWLTLPVPSVYGVFVWSL 117
 QY 129 PAN--AVTLNMLFRTSICITTVYTNLAIADEFLFCVTLPFKIAYHLNGNNWVGEVLCR 186
 DB 118 PLNIMATWVFLKMKVKK-PAVYVMLHLATADLVFVSLPFKISYVPSGSDMQFGSELGR 176
 QY 187 ATTVIFGNMYCSTILLACISINRYLAIVHP---FTYRGIPKHYALVTCLGYNWATVFLY 243
 DB 177 FVTAAFYCNMYASILLMTVISIDRFVAVVPMQSLSWRTIGRASV---TCLAIWALAIAG 233
 QY 244 MLPPFILKQBYLVQPDITTCVDVHNTCCSSSPFQLYYFISLAFGFFGLPFLVLIICYAA 303
 DB 234 VVPLVLKEQITQVGLNITTCVDVHNTCCSSSPFQLYYFISLAFGFFGLPFLVLIICYAA 292
 QY 304 IIRTLN-----AYDHRWLYWYKASLLILVITFCIPAPSNIIILIHANY-YNNNTDGL 355
 DB 293 IIRLSSSAVANRSKGRALF---LSAAVFCIFICFGPTNVLLIAHYSLFSLTSTTEA 349

QY 356 YETIYLIALCLGSLNSCLDPFLYFLMS 381
 DB 350 YFAYLLCVSVSSISSCIDPLIYYAS 375
 RESULT 22
 AAY15082
 ID AAY15082 standard; Protein; 385 AA.
 XX
 AC AAY15082;
 XX
 DT 25-JAN-2000 (first entry)
 XX
 DE Human protease-activated receptor 4 (PAR4).
 XX
 KW Human protease-activated receptor 4; PAR4; thrombin receptor;
 KW phosphoinositide hydrolysis; calcium efflux; platelet aggregation;
 KW platelet activation; thrombin agonist; thrombin antagonist; therapeutic;
 KW wound; blood coagulation; heart attack; stroke; inflammatory response;
 KW proliferative response; atherosclerosis; restenosis; glomerulosclerosis;
 KW pulmonary inflammation; ARDS; adult respiratory distress syndrome.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 48..53
 FT /label= PAR4_activating_peptide
 FT Cleavage-site 47..48
 FT /label= Thrombin_cleavage_site
 XX
 PN WO9943809-A2.
 XX
 PD 02-SEP-1999.
 XX
 PF 11-FEB-1999; 99WO-US02983.
 XX
 PR 27-FEB-1998; 98US-0032397.
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA
 PI Coughlin SR, Kahn M;
 XX
 XX WPI; 1999-619953/53.
 XX
 PT DNA molecules encoding protease-activated receptor 4, useful in
 compound assays for thrombin agonist and antagonist activity -
 XX
 PS Claim 7; Fig 5; 69pp; English.
 XX
 CC The present sequence is a human protease-activated receptor 4
 CC (PAR4), a G protein-coupled thrombin receptor expressed on cell surface.
 CC PAR4 is activated by thrombin and mediates signalling events e.g.
 CC phosphoinositide hydrolysis, calcium efflux and platelet aggregation.
 CC The receptor is highly expressed in spleen cells and likely to be
 CC involved in thrombin-mediated activation of platelets and other
 CC haematopoietic cells. It is used for screening novel thrombin agonists
 CC or antagonists. The agonists are used as therapeutics to treat wounds,
 CC promote clotting and as reagents to activate platelets in diagnostic
 CC tests. Antagonists are used to control blood coagulation, treat heart
 CC attacks and strokes, and block inflammatory and proliferative responses
 CC that occur in normal wound healing and variety of diseases including
 CC atherosclerosis, restenosis, pulmonary inflammation (ARDS) and
 CC glomerulosclerosis.
 XX
 SQ Sequence 385 AA;
 Query Match 24.2%; Score 521.5; DB 20; Length 385;
 Best Local Similarity 33.8%; Pred. No. 1.9e-45;
 Matches 120; Conservative 62; Mismatches 158; Indels 15; Gaps 5;
 QY 29 GLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPNSFEEPP--FSALEGWTGATTVK 86

Db 3 GRLLMLPLVLGSLSGGTQ-----TPSVYDESGSTGGDDSTPSILPAPRGYPCQV---- 53

Qy 87 IKCPESASHLHVKNATMGYLTSSLSKLIPIAYLLVVFVGVPPANAVTLMLFFRTRISIC 146

Db 54 --CANDSDT-LELPDSSRALLLGGWVPRFVLPALYGLVVLVGLPANGALWVLTATQAPRLP 110

Qy 147 TTVFYTNLAIDELFCVTLFPKTAHYHLNGNNWVGEVLCRAITVIFVGNMYCSILLIACI 206

Db 111 STMLMLNLAADULLALPPIAYHLRGORWPFGEAACRAATLAALYGHMYGSLVLLAAV 170

Qy 207 SINRYLAIVHPFTYRGPKHTYALVTCGLVWATVFLVMLPFFILKQBYLYLVOPDITTCDD 266

Db 171 SIORYLALVPLPARALGRRLALGLCMAWLMWMAALALPLTQRTFLRARDVULCHD 230

Qy 267 VHNCTSSPPQLYFYIFSLAFFGFLIPFVLIYYCYAAIIRTNAYDHRWLWYVKASILLI 326

Db 231 ALPLDAQASHWQ-PAFTCALGLGCFLLPALLMCLCYGATLTLAASGRYRGHALRLTAVL 299

Qy 327 VITICFAPSNIILIIHANYNNYNTDGLFYIYIALCLGSLNSCLDPPFLYFLMS 381

Db 290 ASAVAFVPSNLLLLHYSDPSPSAMGNLYGAYVPSLALSTLNSCVDPFYIYVYS 344

RESULT 23

AA50135

ID AA50135 standard; Protein; 385 AA.

AC AA50135;

DT 31-JAN-2000 (first entry)

DE Human protease-activated receptor PAR4.

KW Protease-activated receptor; PAR4; ZCHEMR2; G protein coupled; cellular signalling; protease; tethered ligand; N-terminal; proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist; antagonist; cellular response; physiological response; clotting pathway; platelet; proliferation; differentiation; mediation; inflammatory process; vascular injury; chemotaxis; mitogenesis; growth factor; production.

OS Homo sapiens.

FH Key

FT Region 1..78 Location/Qualifiers

FT Peptide 1..17 /note= "Extracellular N-terminal region"

FT Cleavage-site 17..18 /note= "Signal peptide"

FT Protein 18..385 /note= "Cleaved by signal peptidase"

FT Cleavage-site 47..48 /note= "Mature non-activated human PAR4"

FT Protein 48..385 /note= "Cleaved by thrombin or trypsin to activate"

FT Region 48..53 /note= "Activated human PAR4"

FT (AA50140) Modified-site 56 /note= "Tethered hexapeptide activating ligand"

FT Domain 79..102 /note= "N-glycosylated"

FT Region 103..110 /note= "Transmembrane domain 1"

FT Domain 111..132 /note= "Intracellular loop 1"

FT Region 133..150 /note= "Transmembrane domain 2"

FT Domain 151..172 /note= "Extracellular loop 1"

FT Region 173..191 /note= "Transmembrane domain 3"

FT Domain 192..213 /note= "Intracellular loop 2"

FT Region /note= "Transmembrane domain 4"

FT 214..239

FT /note= "Extracellular loop 2, determines specificity for activating peptide"

FT Region 228..230

FT /note= "These three residues are conserved among PAR1-4"

FT Domain 240..263

FT Region 264..283

FT Domain 284..305

FT Region 306..316

FT Domain 317..343

FT Region 344..385

FT /note= "Intracellular C-terminal region"

XX WO950415-A2.

PN XX

XX 07-OCT-1999.

PF 31-MAR-1999; 99WO-US07100.

XX

PR 01-APR-1998; 98US-0053866.

XX (ZYMO) ZYMOGENETICS INC.

PA (UNIW) UNIV WASHINGTON.

XX

PI Xu W, Presnell SR, Yee DP, Foster DC;

XX WPI; 1999-633640/54.

DR N-PSDB; AA232747, AA232748.

XX

PT Novel protease activated receptor 4, useful for screening for (antagonists for promoting the proliferation and/or differentiation of platelets and in mediating inflammatory events

XX

PS Claim 8; Fig 1; 85pp; English.

XX

CC This sequence represents a human protease-activated receptor, PAR4 (also referred to as ZCHEMR2). The cDNA was identified from EST (expressed sequence tag) sequences with homology to the three known protease-activated receptors (PAR1, PAR2 and PAR3). Protease-activated receptors (PARs) are a subfamily of G protein coupled receptors which are capable of mediating cellular signalling in response to proteases (e.g., thrombin). They are characterised by a tethered peptide ligand at the extracellular N-terminus that is generated by proteolysis. PAR4 is activated by thrombin or trypsin cleavage at Arg47/Gly48, which generates a new N-terminus corresponding to the tethered ligand (a hexapeptide). Agonists of PAR4 are useful for upregulating cellular or physiological responses whereas antagonists are used to downregulate these activities. The PAR4 protein is further useful for dissecting the effects of thrombin or other activating proteases in the clotting pathway from the effects of these proteases at the cellular level. Agonists are specifically useful in promoting the proliferation and/or differentiation of platelets, in mediating inflammatory events, responses to vascular injury, chemotaxis or mitogenesis, and in producing growth factors. Antagonists are useful as research reagents for characterising sites of ligand-receptor interaction.

XX

SQ Sequence 385 AA;

Query Match 24.2%; Score 521.5; DB 20; Length 385;

Best Local Similarity 33.8%; Pred. No. 1.9e-45;

Matches 120; Conservative 62; Mismatches 158; Indels 15; Gaps 5;

Qy 29 GLLLPLPTFCQSGMENDTNNAKPTLPKTFGAPNSFEERP--PSALEGTGATITVK 86

Db 3 GRLLMLPLVLGSLSGGTQ-----TPSVYDESGSTGGDDSTPSILPAPRGYPCQV---- 53

Qy 87 IKCPESASHLHVKNATMGYLTSSLSKLIPIAYLLVVFVGVPPANAVTLMLFFRTRISIC 146

Db 54 --CANDSDT--LELPDSSRALLGWPTRLVPALYGLVVLVGLPANGALWVLAQAPRLP 110
Qy 147 TTVFYTNLATADELFVCVTLPFKIAYHLNGNNWVGEVLCRAATVIFYGNMYCSILLACI 206
Db 111 STMLMNLATADLLALPRLIAYHLRGORWPFGEACRLAATVIFYGNMYCSILLAAV 170
Qy 207 SINRYLAIVHPPTVGRGLPKHTYALVTCGLVWATVFLYMLPFFILKOEYVLQPDITTC 266
Db 171 SLDRYLALVHLPLARALGRRLALGLCMAAWLMAALALPLTLQRTFRLARSDRVLC 230
Qy 267 VHTNCSSPPQLYYFISLAFPGFLIPFVLIYCYAAIIRTNAYDHRWLVYKASLLIL 326
Db 231 ALPLDAQSHWQ--PAFTCLALLGCFLLPALLMCLCYGATLHTLAASGRRYGHALRLAV 289
Qy 327 VITICFAPSNIILIIHHANYNNYNTDGLYFIYLIALCLGSLNSCLDPPFLYFLMS 381
Db 290 ASAVAFVPSNLLLLLHYSDPSAGNLYGAYVPSLALSTLNSCVDPIFYVVS 344

RESULT 24

AA45036
ID AAY45036 standard; Protein; 385 AA.

AC AAY45036;

DT 31-MAY-2000 (first entry)

DE Human protease activated receptor-4.

XX Human; PAR-4; protease activated receptor;
KW antiseptic molecule; PAR antibody; cytostatic; therapeutic;
KW metastatic tumour cell; placental implantation; invasive cell.

OS Homo sapiens.

XX Key Location/Qualifiers
FH Region 43..67

FT /note= "conserved unique region comprising protease
FT activated domain and hirudin binding domain.
FT Therapeutically useful antiseptic molecules are designed
FT based on this region"

FT Region 228..230
FT /note= "conserved residues in second extracellular
FT loop used for designing therapeutically useful
FT antiseptic molecules"

XX WO200008150-A1.

XX 17-FEB-2000.

XX 05-FEB-1999; 99WO-IL00079.

XX 07-AUG-1998; 98IL-0125698.

XX (HADA-) HADASIT MEDICAL RES SERVICES & DEV.

XX Bar-Shavit R;

XX WPI; 2000-205706/18.

XX N-PSDB; AA250775.

XX Treating metastatic tumor cells useful for treating disorders involving
PT placenta implantation in a female comprises administration of an
PT antiseptic molecule complementary to an RNA sequence of a protease
PT activated receptor protein

XX Example 3; Fig 11b; 46pp; English.

XX The patent discloses a method to treat metastatic tumour cells using
CC an antiseptic molecule comprising a polynucleotide complementary to an
CC RNA sequence of a protease activated receptor (PAR) protein, or an
CC antibody capable of binding to a PAR protein. The antiseptic molecules and

CC antibodies of PAR protein are also used to treat disorders associated
CC with implantation of placenta. The present sequence is a human PAR-4
CC protein used for producing antiseptic molecules for treating invasive
CC cells.

XX Sequence 385 AA;

Qy Query Match 24.2%; Score 521.5; DB 21; Length 385;
Best Local Similarity 33.8%; Pred No 1.9e-45;
Matches 120; Conservative 62; Mismatches 158; Indels 15; Gaps 5;

Qy 29 GLLLLPTFCQSGMENDTNLAKPLPIKTFPGAPPNSFEFP--FSALEGWGTATITVK 86
Db 3 GRLLMLPLVGLSGLGTO-----TPSVYDESGTGGDDSTESILPAPRGYPGVQV--- 53

Qy 87 IKCPESASHLVKVNATMGYLTSSLSSTKLIPALYLLVGVVGPANAVTLMWLFRTSIC 146

Db 54 --CANDSDT--LELPDSSRALLGWPTRLVPALYGLVVLVGLPANGALWVLAQAPRLP 110

Qy 147 TTVFYTNLATADELFVCVTLPFKIAYHLNGNNWVGEVLCRAATVIFYGNMYCSILLACI 206

Db 111 STMLMNLATADLLALPRLIAYHLRGORWPFGEACRLAATVIFYGNMYCSILLAAV 170

Qy 207 SINRYLAIVHPPTVGRGLPKHTYALVTCGLVWATVFLYMLPFFILKOEYVLQPDITTC 266

Db 171 SLDRYLALVHLPLARALGRRLALGLCMAAWLMAALALPLTLQRTFRLARSDRVLC 230

Qy 267 VHTNCSSPPQLYYFISLAFPGFLIPFVLIYCYAAIIRTNAYDHRWLVYKASLLIL 326

Db 231 ALPLDAQSHWQ--PAFTCLALLGCFLLPALLMCLCYGATLHTLAASGRRYGHALRLAV 289

Qy 327 VITICFAPSNIILIIHHANYNNYNTDGLYFIYLIALCLGSLNSCLDPPFLYFLMS 381

Db 290 ASAVAFVPSNLLLLLHYSDPSAGNLYGAYVPSLALSTLNSCVDPIFYVVS 344

RESULT 25

AA47623

ID AAB47623 standard; Protein; 385 AA.

XX AAB47623;

XX 21-JAN-2002 (first entry)

XX Human PAR4.

XX Human; protease-activated receptor; PAR4; N-terminal; activated PAR4;
KW platelet aggregation; inhibition; tumour cell; proliferation.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 48..53

FT /label= N-terminal of activated PAR4

XX WO200158930-A1.

XX 16-AUG-2001.

XX 06-FEB-2001; 2001WO-US03807.

XX 09-FEB-2000; 2000US-0500646.

XX (ZYMO) ZYMOGENETICS INC.

XX Baidur N, West RR;

XX WPI; 2001-656678/75.

XX N-PSDB; AA43632.

XX Peptides comprising an amino acid sequence are capable of stimulating
PT protease-activated receptor 4 which are useful in diagnosis and therapy
PT e.g. inhibiting tumor cell proliferation and stimulating platelet

PT aggregation -
XX -PS Disclosure; Page 58-60; 84pp; English.
XX
CC This sequence shows human protease-activated receptor, PAR4. Peptides
CC derived from, or based on, the N-terminal of activated PAR4 (see
CC AAB47624-77), are capable of stimulating PAR4. These peptides may be
CC used to activate PAR4 at lower concentrations than wild type PAR4.
CC These peptides may be used to stimulate platelet aggregation, and for
CC inhibiting tumor cell proliferation.
XX
SQ Sequence 385 AA;

Query Match 24.2%; Score 521.5; DB 22; Length 385;
Best Local Similarity 33.8%; Pred. No. 1.9e-45;
Matches 120; Conservative 62; Mismatches 158; Indels 15; Gaps 5;

QY 29 GLILLPTFCQSGMENDNNLAKPTLPKTRGAPNSPEFP--FSALEGTGATITVK 86
Db 3 GRULLPLVLGSLSGTQ-----TPSVDESSTGGDDSTPSILPAPRGYQGV---- 53

QY 87 IKCPESASHLVKNATMGVLTSSLSKLPATVLLVVGVPANAVTLMLEFTRRSIC 146
Db 54 --CANDSDT-LELPDSSRALLGWPTRLVPALYGLVVLGVLPAAGLAWLATQAPRLP 110

QY 147 TVFYVNLATADFLFCVTLPFKIAYHLNGNWNVGEVLCRATTVI FYGNMYCSILLACI 206
Db 111 STMLLANLATADLLALPPRIAYHLRGQRWPFGEACRLATAALYGHMYGSVLLAAV 170

QY 207 SINRYLAIVHPFTYRGLPKITYALVTCGLVWATVFLMPLFFKQBYVLVQDPITTC 266
Db 171 SLDRYLAIVHLPRARALGRRLALGLCMAALMAALALPLTLQRTFLRSLDRVLCHD 230

QY 267 VHTNCSSPFQLYFYFISLAFFGLIPFVLLIYCYAIIHTLNAYDHRMLWYKASILL 326
Db 231 ALPDAQASHWQ-PATFCLALLGCFLLPALLMLLCYGTATHTLAASGRRYGHARLFAVL 289

QY 327 VIPTICFAPSNIILIHANYNNYNTDGLYFIYLIALLCLGSLNSCLDPPFLYFLMS 381
Db 290 ASAVAFVPSNLLLLHYSPSPSAMGNLYGAVPSLALSTLNSCVDFFIYYVVS 344

RESULT 26
AAY50138
ID AAY50138 standard; Protein; 385 AA.
XX
AC AAY50138;
XX
DT 31-JAN-2000 (first entry)
XX
DE Mutant human protease-activated receptor PAR4, R47A.
XX
KW Protease-activated receptor; PAR4; G protein coupled;
KW cellular signalling; protease; tethered ligand; N-terminal;
KW proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist;
KW antagonist; cellular response; physiological response; clotting pathway;
KW platelet; proliferation; differentiation; mediation;
KW inflammatory process; vascular injury; chemotaxis; mitogenesis;
KW growth factor; production; mutant.
XX
OS Homo sapiens.
OS Synthetic.
XX
PH Key Location/Qualifiers
FT Region 1..78 /note= "Extracellular N-terminal region"
FT Peptide 1..17 /note= "Signal peptide"
FT Cleavage-site 17..18 /note= "Cleaved by signal peptidase"
FT Protein 18..385 /note= "Mature non-activated human PAR4"
FT Region 48..53

/note= "Tethered hexapeptide activating ligand
(AAY50140)" Modified-site 56
/note= "N-glycosylated"
79..102
Domain
/note= "Transmembrane domain 1"
103..110
Region
/note= "Intracellular loop 1"
111..132
Domain
/note= "Transmembrane domain 2"
133..150
Region
/note= "Extracellular loop 1"
151..172
Domain
/note= "Transmembrane domain 3"
173..191
Region
/note= "Intracellular loop 2"
192..213
Domain
/note= "Transmembrane domain 4"
214..239
Region
/note= "Extracellular loop 2, determines specificity for
activating peptide"
228..230
Region
/note= "These three residues are conserved among PAR1-4"
240..263
Domain
/note= "Transmembrane domain 5"
264..283
Region
/note= "Intracellular loop 3"
284..305
Domain
/note= "Transmembrane domain 6"
306..316
Region
/note= "Extracellular loop 3"
317..343
Domain
/note= "Transmembrane domain 7"
344..385
Region
/note= "Intracellular C-terminal region"
WO9950415-A2.
07-OCT-1999.
31-MAR-1999; 99WO-US07100.
01-APR-1998; 98US-0053866.
(ZYMO) ZYMOGENETICS INC.
(UNIW) UNIV WASHINGTON.
Xu W, Presnell SR, Yee DP, Foster DC;
WPI; 1999-633640/54.
Novel protease activated receptor 4, useful for screening for
(ant)agonists for promoting the proliferation and/or differentiation of
platelets and in mediating inflammatory events -
Example 2; Page -; 85pp; English.
This sequence represents a mutant human protease-activated receptor PAR4
R47A, compared with PAR4 mutant R68A (AAY50139) in studies to evaluate
the importance of a protease cleavage site between residues 47 and 48.
This mutant PAR4 was unable to be cleaved by thrombin or trypsin, in
contrast to the wild-type (AAY50135) or the R68A mutant. This cleavage
activates PAR4, exposing a tethered hexapeptide ligand at the new
N-terminus. Protease-activated receptors (PARs) are a subfamily of G
protein coupled receptors which are capable of mediating cellular
signalling in response to proteases. Agonists of PAR4 are useful for
upregulating cellular or physiological responses whereas antagonists
are used to downregulate these activities. The PAR4 protein is
further useful for dissecting the effects of thrombin or other
activating proteases in the clotting pathway from the effects of these
proteases at the cellular level. Agonists are specifically useful in
promoting the proliferation and/or differentiation of platelets, in
mediating inflammatory events, responses to vascular injury, chemotaxis
or mitogenesis, and in producing growth factors. Antagonists are useful

XX Disclosure; Fig 14; 118pp; English.
 XX
 CC The sequence is that of the human thrombin receptor fragment.
 CC This is used to exemplify a new expression vector for producing
 CC heterologous polypeptides in a halobacterial host.
 CC See also AAR60891-9.
 XX
 SQ Sequence 425 AA;
 Query Match 23.9%; Score 516.5; DB 15; Length 425;
 Best Local Similarity 32.6%; Pred. No. 7.2e-45;
 Matches 123; Conservative 77; Mismatches 144; Indels 33; Gaps 11;
 QY 30 LLLLLPTFCQ-----SGMENDTNLAKPTLPKTRGAPPN-SPEEPFSALEGTGA 81
 Db 1 MLELLPTAVEGVSAQIQALDYKDDDDVDATLDRSFLLRPNNDKYEPFWEDEKNESGL 60
 QY 82 T-----ITVKIKPEESASHLVKQVATGKLTSSLSKLPAIYLLVFGVGVAN--AVTL 135
 Db 61 TEYRLVSINKSPLOKQPAFISEDASGLTSSWLTFLFVSVYTGTVVSLPNIIMAVV 120
 QY 136 WMLFFTRISCTTVFVFNLAIAADFLFCVTLPPKIAVHLNGNNVFGVLCRATTIVFGN 195
 Db 121 FILKMKVKK-PAVYMLHLATADVLSVLPFKISYFGSDMQFGSELRCFVTAAYCYN 179
 QY 196 MYCSILLACISINRYLAIVHP---FYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQ 252
 Db 180 MYASILLMTVISIDFLAVVPMQSLSWRTLGRSF---TCLAIWALAIAGVPLVLKQ 236
 QY 253 EYVLQPDITTDCHDVNTCESSPFPQYFISLAFFGLIPFVLIYCYAAIIRTLN--- 309
 Db 237 TIQVPLNITTDCHDVNTLLEG-YYAYYFSAFSAVFFVPLIISTVCYVSIIRCLSSA 295
 QY 310 ----AYDHRWLWVVKASLLILVITFCFAPSNIILIIHHANY-YNNITDGLFYLIALC 364
 Db 296 VANRSKKSRALF---LSAAVFCIFICFGPTNLLIIAHYSFUSHTSTTEAAYFALLVCV 352
 QY 365 LGSLSNCLDPFLYFLMS 381
 Db 353 VSSISSCIDPLIYYAS 369
 RESULT 29
 ID AAW16314
 AC AAW16314 standard; Protein; 892 AA.
 XX
 DT 16-AUG-1997 (first entry)
 DE Human thrombin receptor-yeast G-alpha protein fusion.
 DE
 KW G-protein coupled receptor; agonist; antagonist; assay;
 KW G-alpha protein; Gpalp; GPAL gene; thrombin receptor.
 XX
 OS Chimaeric Homo sapiens;
 OS Chimaeric Saccharomyces cerevisiae.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..13
 FT /label= Sig_peptide
 FT /note= "alpha-factor (STE2) signal peptide"
 FT Protein 14..892
 FT /label= ThR-GPAL_fusion
 FT Region 14..417
 FT /label= ThR
 FT /note= "thrombin receptor amino acids 22-425"
 FT Peptide 418..420
 FT /label= Linker
 FT Region 421..892
 FT /label= G-alpha_protein
 XX

PN WO9711159-A1.
 XX
 PD 27-MAR-1997.
 XX
 PF 20-SEP-1996; 96WO-US15203.
 XX
 PR 20-SEP-1995; 95US-0004023.
 XX
 PA (HEAR-) HEARTLAND BIOTECHNOLOGIES LLC.
 XX
 PI Boulton TG, Das PO, Mandell RB, McMullen TW;
 XX
 DR WPI; 1997-202868/18.
 XX
 DR N-PSDB; AAT62461.
 XX
 PT New yeast constructs for use in screening assays - which express a
 PT fusion comprising a mammalian or fungal transmembrane receptor and
 PT a yeast G-alpha protein
 XX
 PS Example 2; Fig 12A-G; 109pp; English.
 XX
 CC A fusion protein (AAW16314) comprises the human thrombin receptor
 CC (amino acids 22 to C-terminal threonine) covalently linked to
 CC yeast G-alpha protein Gpalp (encoded by the GPAL gene). It is
 CC expressed by a ThR-GPAL gene fusion (AAT62461) in vector pRMHT16.
 CC Thrombin-dependent activation of the yeast mating pathway by the
 CC ThR-GPAL fusion protein was observed in ste2 gpal- yeast cells.
 CC The polypeptide is an example of novel fusion proteins between
 CC eukaryotic heterotrimeric G-protein coupled receptors and yeast
 CC G-alpha proteins that can be coupled to the pheromone-induced
 CC signal transduction pathway of yeast and used in screening assays
 CC to identify agonists or antagonists of the receptor.
 XX
 SQ Sequence 892 AA;
 Query Match 23.7%; Score 512; DB 18; Length 892;
 Best Local Similarity 34.0%; Pred. No. 5.5e-44;
 Matches 118; Conservative 71; Mismatches 132; Indels 26; Gaps 10;
 QY 53 TLPIKTERGAPPN-SPEEPFSALEGTGAT---ITVKIKPEESASHLVKQVATGKLT 107
 Db 29 TLDRSFLLRPNNDKYEPFWEDEKNESGLTYRLVSINKSPLOKQPAFISEDASGL 88
 QY 108 TSSLSTKLIPAIYLLVFGVGVAN--AVTLWMLFFTRISCTTVFVFNLAIAADFLFCVTL 165
 Db 89 TSSWLTFLFVPSYTGTVVSLPNIIMAVIWFILKMKVKK-PAVYMLHLATADVLSV 147
 QY 166 PKIAYHLNGNNVFGVLCRATTIVFYGNMYCSILLACISINRYLAIVHP---FTYRG 222
 Db 148 PFKISYFSGSDMQFGSELRCFVTAAYCYNMYASILLMTVISIDRFLAVVPMQSLSWRT 207
 QY 223 LPKHTYALVTCGLVWATVFLYMLPFFILKQYVLPQDITTDCHDVNTCESSPFPQYF 282
 Db 208 LGRASF---TCLAIWALAIAGVPLVLKQETIQVPLNITTDCHDVNTLLEG-YIAYYF 263
 QY 283 ISLAFFGLIPFVLIYCYAAIIRTLN-----AYDHRWLWVVKASLLILVITFCFAP 335
 Db 264 SAFAVFFVPLIISTVCYVSIIRCLSSSAVANRSKKSRALF---LSAAVFCIFICFGP 320
 QY 336 SNIIIIHHANY-YNNITDGLFYLIYALCLGSLNSCLDPFLYFLMS 381
 Db 321 TNVLLIIAHYSFUSHTSTTEAAYFALLVCVSSISSCIDPLIYYAS 367
 RESULT 30
 ID ABG35298
 XX ABG35298 standard; Protein; 402 AA.
 AC ABG35298;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human PAR1 type thrombin receptor delta 1-49TR.

XX Human; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW inflammatory disease; cell proliferative disease.
XX Homo sapiens.

OS

XX JP2002010784-A.

XX 15-JAN-2002.

XX 29-JUN-2000; 2000JP-0196514.

XX 29-JUN-2000; 2000JP-0196514.

XX (TEIJ) TEIJIN LTD.

XX WPI; 2002-321520/36.

DR N-PSDB; ABK70887.

XX An inhibitor of cell growth mediated by thrombin used to treat
PT inflammatory and cell proliferative diseases -

XX Disclosure; Page 22-24; 44pp; Japanese.

XX The invention relates to a polypeptide or a compound which can inhibit
CC cell growth caused by thrombin. The polypeptide/compound combines to a
CC specific region of the structure of PAR1 type human thrombin receptor
CC participating to cell growth. Preferably, the compound contains the
CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
CC amino acid or peptide sequence). Also included are a modified PAR1 type
CC thrombin receptor gene or its fragment used for obtaining the above
CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a PAR1 type thrombin receptor (or a modified
CC version).

XX Sequence 402 AA;

Query Match 23.5%; Score 507.5; DB 23; Length 402;
Best Local Similarity 33.0%; Pred. No. 5.7e-44;
Matches 121; Conservative 71; Mismatches 138; Indels 37; Gaps 11;

QY 32 LLLPTFCQSGMENDNNLAKPTLPIKTFRCAPNSPEEPFSALEGTGAT----ITVKI 87
DB 6 LLLVAACFS-----LCGPLLSART-----RADKYEPPFWEDEKNEGLTEYRLVSINK 53

QY 88 KCPEESASHLRVKNATGYLTSSLTSLKLIPIAIVLLVFGVGPAN--AVTLWMLFFRTRSI 145
DB 54 SSPLQKLPAPFISEDASGYLTSSWTLTFVPSVTGVFVSLPLNINAIIVFILKMKVKK- 112

QY 146 CTTVFYNNLAIADFLCVTLPPKIAHYLNGNNWVGEVLCRATTVIFYGNMYCSILLAC 205
DB 113 PAVYMLHLATADLVFVSLPFIKISYFSGSDQFSGELCRFVTAAYFCNMYASILLMTV 172

QY 206 ISINRYLAIVHP---FTYRGLPKHTALVTCGLVWATVFLYMLPFFILKQEVYLVQPDIT 262
DB 173 ISIDRFVAVYPMQSLWSRITLGRASP---TCLAIWALAIAGVVPVLLKQTIQVPGNLT 229

QY 263 TCHDVHNTCESSPPQLYIYISIAFFGLFIPFVLLIYCYAAIIRTNL-----AYDHRW 315
DB 230 TCHDVNLNETLEG-YIAYIYSAFSAVFFVPLIISTVCYVSIIRCLSSSAVANRKSRA 288

QY 316 LWYKASLLILVPTICFAPSNIILIIHANY-YNNNTDGLFIYLIACLSGNSCLDP 374
DB 289 LF---LSAAVFCIFIIICFGETNVLIIAHYFSLSTTAAAYFAYLLVCVSSISCCIDP 345

QY 375 FLYFLMS 381

DB 346 LIYYAS 352

RESULT 31

ABG35299

XX ABG35299 standard; Protein; 371 AA.

XX AC ABG35299;

XX 15-JUL-2002 (first entry)

XX Human PAR1 type thrombin receptor delta 1-80.

XX Human; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW inflammatory disease; cell proliferative disease.

XX Homo sapiens.

XX JP2002010784-A.

XX 15-JAN-2002.

XX 29-JUN-2000; 2000JP-0196514.

XX 29-JUN-2000; 2000JP-0196514.

XX (TEIJ) TEIJIN LTD.

XX WPI; 2002-321520/36.

DR N-PSDB; ABK70888.

XX An inhibitor of cell growth mediated by thrombin used to treat
PT inflammatory and cell proliferative diseases -

XX Disclosure; Page 24-25; 44pp; Japanese.

XX The invention relates to a polypeptide or a compound which can inhibit
CC cell growth caused by thrombin. The polypeptide/compound combines to a
CC specific region of the structure of PAR1 type human thrombin receptor
CC participating to cell growth. Preferably, the compound contains the
CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
CC amino acid or peptide sequence). Also included are a modified PAR1 type
CC thrombin receptor gene or its fragment used for obtaining the above
CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a PAR1 type thrombin receptor (or a modified
CC version).

XX Sequence 371 AA;

Query Match 23.2%; Score 499.5; DB 23; Length 371;
Best Local Similarity 36.9%; Pred. No. 3.5e-43;
Matches 107; Conservative 61; Mismatches 101; Indels 21; Gaps 8;

QY 105 GYLTSLSLTKLIPIAIVLLVFGVGPAN--AVTLWMLFFRTRISICTTVFVNNLAIAADFLPC 162
DB 40 GYLTSLSLTKLIPIAIVLLVFGVGPAN--AVTLWMLFFRTRISICTTVFVNNLAIAADFLPC 98

QY 163 VTLPPKIAHYLNGNNWVGEVLCRATTVIFYGNMYCSILLACISINRYLAIVHP---FT 219
DB 99 SVLPFKISYFSGSDQFSGELCRFVTAAYFCNMYASILLMTVISIDRFVAVYPMQSL 158

QY 220 YRGLPKHTALVTCGLVWATVFLYMLPFFILKQEVYLVQPDITTCCHDVHNTCESSPPQ 279
DB 159 WRTLGRASP---TCLAIWALAIAGVVPVLLKQTIQVPGNLTTCCHDVHNTCESSPPQ 214

QY 280 YFISIAFFGLFIPFVLLIYCYAAIIRTNL-----AYDHRWLVWYKASLLILVPTIC 332
DB 215 YFISIAFFGLFIPFVLLIYCYAAIIRTNL-----AYDHRWLVWYKASLLILVPTIC 271

QY 333 FAPSNIILIIHANY-YNNNTDGLFIYLIACLSGNSCLDPFLYFLMS 381

DB 272 FGPTNVLIIAHYFSLSTTAAAYFAYLLVCVSSISCCIDPLIYYAS 321

Db 70 LMLASVLPQIYHCHNRHHVFGVLLCNVTVAFYANMYSSILMTWCISVERFLGVLYPL 129
 QY 219 TYRGLPKHTYALVTCGLVWATVFLYMLPFILKQEVYLVQPDITTCCHDVHN-TCESSSPF 277
 Db 130 SSKRWRRRYAAVACAGTWLLLTALSPARTDLYTPVHALGIITCFDLKWTMLPSVAM 189
 QY 278 QLYYFSLAPFGFLPELVLIYCAA----IIRTLNAYD-HRWLVWVKASLLILVIFTIC 332
 Db 190 WAVFLTFIFLLFIPFVITVACYTATILKLTTEAHGREQSAAGVLAUVLLAFVTC 249
 QY 333 FAPSNIIILIH-HANYYYNNTDGLYFIYLIALCLGSLNSCLDPLFYFLMSK 382
 Db 250 FAPNNEVLLAHIVSRIFYGKS--YYHVYKLTCLCLNCLDPLFYVYFASR 298

RESULT 37
 AAG78530
 ID AAG78530 standard; Protein; 359 AA.
 XX
 AC AAG78530;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE G-protein coupled receptor (designated Paul) amino acid sequence.
 XX
 KW G-protein coupled receptor; paul; 7TM receptor; GPC receptor;
 KW human; antibacterial; virucide; fungicide; protozoacide;
 KW anti-Human immunodeficiency Virus; analgesic; cytosstatic; antidiabetic;
 KW anorectic; antiaesthematic; antiparkinsonian; cardiant; hypertensive;
 KW hypotensive; osteopathic; antianigmal; antimanic; cerebroprotective;
 KW antiulcer; antiallergic; antidepressant; antimigraine; antiemetic;
 KW tranquiliser; neuroleptic; neuroprotective; nootropic; anticonvulsant;
 KW Human immunodeficiency virus type 1; HIV-2; HIV-1; HIV-2; pain; cancer;
 KW diabetes; obesity; anorexia; bulimia; asthma; Parkinson's disease;
 KW acute heart failure; hypertension; myocardial infarction; urinary retention;
 KW osteoporosis; angina pectoris; myocardi infarction; stroke; ulcer;
 KW allergy; benign prostatic hypertrophy; migraine; vomiting; anxiety;
 KW schizophrenia; manic depression; delirium; dementia; mental retardation;
 KW dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome.
 XX
 OS Homo sapiens.
 XX
 PN US2001029032-A1.
 XX
 PD 11-OCT-2001.
 XX
 PF 18-DEC-2000; 2000US-0739151.
 XX
 PR 09-OCT-1998; 98US-103789P.
 PR 06-OCT-1999; 99US-0413534.
 XX
 PA (ZHU)/ ZHU Y.
 PA (LIXX)/ LI X.
 PA (VAVT)/ VAWTER L.
 PI
 PI Zhu Y, Li X, Vawter L;
 XX
 DR WPI; 2001-647985/74.
 DR N-PSDS; AAI64231.
 XX
 PT New G-protein coupled receptor polypeptide, referred as Paul and
 PT encoding polynucleotide, useful for diagnosing and treating cancers,
 PT infections, neurological disorders, diabetes, asthma and identifying
 PT modulators -
 XX
 XX Claim 1c; Page 13-14; 15pp; English.
 XX
 CC The invention relates to an isolated polypeptide, a member of G-protein
 CC coupled receptor family of polypeptides, comprising a fully defined
 CC sequence of 359 amino acids, its 95% identical sequence, or a polypeptide
 CC encoded by a polynucleotide comprising a fully defined sequence of 1080
 CC base pairs defined in the specification, or a fragment or variant of it.

CC The activity of the protein of the invention may be described as
 CC antibacterial, virucide, fungicide, protozoacide, anti-Human
 CC immunodeficiency Virus, analgesic, cytosstatic, antidiabetic, anorectic,
 CC antiaesthematic, antiparkinsonian, cardiant, hypertensive, hypotensive,
 CC osteopathic, antianigmal, antimanic, cerebroprotective, antiulcer,
 CC antiallergic, antidepressant, antimigraine, antiemetic, tranquiliser,
 CC neuroleptic, neuroprotective, nootropic and anticonvulsant. Polypeptides
 CC and polynucleotides of the invention are useful in diagnosis and in
 CC identifying compounds such as agonists and antagonists which are useful
 CC in therapy. They are also useful for treating diseases, including
 CC infections such as bacterial, fungal, protozoan and viral infections,
 CC particularly infections caused by Human immunodeficiency virus type 1 or
 CC 2 (HIV-1 or HIV-2), pain, cancers, diabetes, obesity, anorexia, bulimia,
 CC asthma, Parkinson's disease, acute heart failure, hypotension,
 CC hypertension, urinary retention, osteoporosis, angina pectoris,
 CC myocardial infarction, stroke, ulcers, allergies, benign prostatic
 CC hypertrophy, migraine, vomiting, psychotic and neurological disorders,
 CC including anxiety, schizophrenia, manic depression, depression, delirium,
 CC dementia, and severe mental retardation, dyskinesias such as Huntington's
 CC disease or Gilles de la Tourette's syndrome. Polypeptides and
 CC polynucleotides of the invention may also be useful as diagnostic
 CC reagents, for example in the detection of mutations in the associated
 CC gene, chromosome localisation studies and expression pattern
 CC determination. The current sequence represents a G-protein coupled
 CC receptor (designated "Paul" in the specification) amino acid sequence.
 XX
 SQ Sequence 359 AA;
 Query Match 21.3%; Score 460; DB 22; Length 359;
 Best Local Similarity 34.4%; Pred. No. 4.2e-39;
 Matches 100; Conservative 54; Mismatches 125; Indels 12; Gaps 7;
 QY 101 NATMGVLTSSSLTKLIPALYLLAVFVGVPPANAVTLMLEFR--TBSICTTVFVTNLAID 158
 Db 11 NATQLWRNPATAVALPVVYSLVAAVSPGNLFLSLVLCRRMGPRS-PSVIFMINLSVTD 69
 QY 159 FLFCVTLPPKIAHYHNGNNWFEVLCRATTVFYGNMYCSILLACISINRYLAIVHPF 218
 Db 70 LMLASVLPQIYHCHNRHHVFGVLLCNVTVAFYANMYSSILMTWCISVERFLGVLYPL 129
 QY 219 TYRGLPKHTYALVTCGLVWATVFLYMLPFILKQEVYLVQPDITTCCHDVHN-TCESSSPF 277
 Db 130 SSKRWRRRYAAVACAGTWLLLTALSPARTDLYTPVHALGIITCFDLKWTMLPSVAM 189
 QY 278 QLYYFSLAPFGFLPELVLIYCAA----IIRTLNAYD-HRWLVWVKASLLILVIFTIC 332
 Db 190 WAVFLTFIFLLFIPFVITVACYTATILKLTTEAHGREQSAAGVLAUVLLAFVTC 249
 QY 333 FAPSNIIILIH-HANYYYNNTDGLYFIYLIALCLGSLNSCLDPLFYFLMSK 382
 Db 250 FAPNNEVLLAHIVSRIFYGKS--YYHVYKLTCLCLNCLDPLFYVYFASR 298

RESULT 38
 AAG80966
 ID AAG80966 standard; Protein; 359 AA.
 XX
 AC AAG80966;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE Human, nGPCR5 #2.
 XX
 KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disorder;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder/attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective.

OS Homo sapiens.
 XX WO200136473-A2.
 XX 25-MAY-2001.
 XX 16-NOV-2000; 2000WO-US31581.
 XX 16-NOV-1999; 99US-0165838.
 XX 17-NOV-1999; 99US-0166071.
 XX 19-NOV-1999; 99US-0166678.
 XX 28-DEC-1999; 99US-0173396.
 XX 22-FEB-2000; 2000US-0184129.
 XX 28-FEB-2000; 2000US-0185421.
 XX 28-FEB-2000; 2000US-0185554.
 XX 02-MAR-2000; 2000US-0186530.
 XX 03-MAR-2000; 2000US-0186811.
 XX 09-MAR-2000; 2000US-0188114.
 XX 17-MAR-2000; 2000US-0190310.
 XX 21-MAR-2000; 2000US-0190800.
 XX 20-APR-2000; 2000US-0198568.
 XX 02-MAY-2000; 2000US-0201190.
 XX 08-MAY-2000; 2000US-0203111.
 XX 25-MAY-2000; 2000US-0207094.
 XX (PHAA) PHARMACIA & UPJOHN CO.
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX WPI; 2001-389826/41.
 DR N-PSDB; AAH51006.
 XX New G protein-coupled receptor (NGPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX Claim 37; Page 88; 261pp; English.
 XX The present invention relates to novel G protein-coupled receptors
 CC (NGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
 CC sequence is one such G protein-coupled receptor. GPCRs are also known as
 CC seven transmembrane receptors and function in signal transduction. The
 CC NGPCRx coding sequences are useful for screening a human to diagnose a
 CC disorder affecting the brain or a genetic predisposition, specifically
 CC schizophrenia. NGPCRx are useful for identifying compounds useful for
 CC treating schizophrenia. Detection of NGPCRx in a sample is useful as a
 CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
 CC failure, rheumatoid arthritis, CNS disorders, infections such as HIV-1,
 CC metabolic and cardiovascular diseases, proliferative disorders and
 CC hormonal disorders. Modulators of NGPCRx activity have the utility for
 CC treating neurological disorders, including schizophrenia, ADHD/ADD
 CC (attention deficit-hyperactivity disorder/attention deficit disorder),
 CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
 CC migraine and senile dementia. Additional disorders include inflammatory
 CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
 CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
 CC diseases e.g. inflammatory bowel disease.
 XX Sequence 359 AA;
 SQ Query Match 21.3%; Score 460; DB 22; Length 359;
 Best Local Similarity 34.4%; Pred. No. 4.2e-39;
 Matches 100; Conservative 54; Mismatches 125; Indels 12; Gaps 7;
 QY 101 NATGTYLTSLSKLPALVLLVFGVGPANAVTLMWLFRR--TRSICTTFTVFNIAIAD 158
 DB 11 NATLQMLNPAIAVALPWTSLVAASIPGNLFSWLVCRRMGPRS-PSVIFMNLSTVD 69
 QY 159 FLECVTLFPKVIATHLGNVNVFGEVLCRATVIFYGNWYCSILLACISINRVAIVHPF 218
 DB 70 LMLASVLPFQIYYHCRHHRVFGVLLCNVTVTAFYANNWSSILMTTCISVERFLGVLYPL 129

QY 219 TYRGLPKXTYALVTGCLVWATVFLVWLPFFILKQEYVLVQPDITTTCHDVHN--TCSSSSPF 277
 DB 130 SSKRWRRRRYAAACAGTGLLLLTALSPILARTDLYPVHALGIITCFDVLKWTMLPSVAM 189
 QY 278 QLYYFISLAFFGFLIPFVLIIYCYAA---IIRTLNAYD-HRWLWVVKASLLILVIFETIC 332
 DB 190 WAVEFLFTIFILLFLIPFVITVACYTATIKLURTEAHGREQRRRAVGLAAVVLLAFVTC 249
 QY 333 FAPSNILIIH-HANVYYNNNTDGLYFIYLIALCLASLNSCLDPFFLYFLMSK 382
 DB 250 FAPNNEVLLAHIVSRIFYGKS--YHYVVKLTCLCLNCLNCLDPFFVYFASR 298
 RESULT 39
 AAB62285
 ID AAB62285 standard; Protein; 359 AA.
 XX AAB62285;
 AC AAB62285;
 XX 29-JUN-2001 (first entry)
 XX Human G-protein coupled receptor, PAUL.
 KW G-coupled protein receptor; PAUL; anti-HIV; antibacterial; antiviral;
 KW antifungal; protozoacide; cytostatic; antidiabetic; anorectic; human;
 KW antiasthmatic; antiparkinsonian; cardiac; hypertensive; osteopathic;
 KW antianginal; cerebroprotective; antiulcer; antimigraine; antiemetic;
 KW tranquilizer; nootropic; anticonvulsant; neuroleptic; vaccine.
 XX Homo sapiens.
 OS WO200125280-A1.
 XX 12-APR-2001.
 XX 03-OCT-2000; 2000WO-US27228.
 XX 06-OCT-1999; 99US-0413534.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX Zhu Y, Li X, Vawter L;
 PI WPI; 2001-273570/28.
 DR N-PSDB; AAF57649.
 XX Novel G-coupled protein receptor, PAUL useful for treating diseases
 PT such as microbial infections, cancers, obesity, asthma, diabetes,
 PT hypotension, osteoporosis, myocardial infarction, stroke, ulcer and
 PT allergy -
 XX Claim 1; Page 25; 34pp; English.
 XX This represents a human G-coupled protein receptor (designated PAUL).
 CC The PAUL polypeptide and polynucleotide are useful for treating and
 CC diagnosing infections such as bacterial, fungal, protozoan and viral
 CC infections, particularly infections caused by human immunodeficiency
 CC virus (HIV)-1 or HIV-2, pain, cancers, diabetes, obesity, anorexia,
 CC bulimia, asthma, Parkinson's disease, acute heart failure, hypotension,
 CC hypertension, urinary retention, osteoporosis, angina pectoris, stroke,
 CC myocardial infarction, ulcers, allergies, benign prostatic hypertrophy,
 CC migraine, vomiting, psychotic and neurological disorders including
 CC anxiety, schizophrenia, manic depression, depression, delirium, dementia
 CC and severe mental retardation, and dyskinesia such as Huntington's
 CC disease or Gilles de la Tourette's syndrome. The PAUL sequences are
 CC useful as vaccines to induce an immunological response in a mammal. The
 CC PAUL polypeptide is also useful for identifying its modulators. The PAUL
 CC polynucleotide is valuable for chromosome localization studies and for
 CC tissue expression studies.
 XX Sequence 359 AA;
 SQ Query Match 21.3%; Score 460; DB 22; Length 359;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: June 29, 2003, 07:50:54 ; Search time 1208.52 Seconds
(without alignments)
5454.265 Million cell updates/sec

Title: US-09-208-629f-3
Perfect score: 2136
Sequence: 1 TLYTQHPVAGSQDIKMKIL.....AMARPLXRRPRDIWIHAW 407

Scoring table:
BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlh
-O=/cgn2_1/USPTO_spool/US09208629/runat_24062003_101625_12930/app_query.fasta_1.1166
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629 @CGN 1.1 2013 @runat_24062003_101625_12930 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_othr:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	45.2	965	585	17	A2820748
2	42.6	911	529	9	AA177828
3	36.6	782.5	581	13	BM258063
4	31.4	670	554	13	BM255484
5	22.4	479.5	839	14	BQ897419
6	21.8	466.5	821	14	BM946931
7	21.6	462	938	17	CNS04SUP
8	20.8	445	381	10	BB842550
9	20.8	443.5	765	13	BI183645
10	19.5	416	801	13	BG924078
11	19.5	415.5	624	13	BI394593
12	19.0	406	685	9	AL660446
13	18.4	392.5	808	9	AJ456719
14	18.3	391.5	741	9	AJ453402
15	18.2	389.5	800	9	AJ452824
16	18.2	388	845	9	AJ456135
17	18.1	386	904	9	AL547762
18	18.1	386	931	9	AL547762
19	18.0	384	860	9	AJ446632
20	17.9	383	602	13	BM439733
21	17.9	382.5	760	9	AJ456491
22	17.9	382.5	773	9	AJ451396
23	17.9	382.5	890	9	AJ456561
24	17.9	381.5	747	9	AJ446611
25	17.8	381	946	9	AL551903
26	17.7	377.5	2020	11	BC013202
27	17.7	377.5	2542	11	AK017378
28	17.7	377.5	3001	11	AK005013
29	17.5	374.5	730	9	AJ455912
30	17.5	374.5	740	9	AJ450921
31	17.5	373.5	749	9	AJ447002
32	17.4	371.5	766	9	AJ450835
33	17.2	366.5	746	9	AJ450259
34	17.0	363.5	942	14	BQ896389
35	16.9	360.5	945	9	AL713459
36	16.9	360.5	798	9	AJ451155
37	16.7	356	542	9	AI942392
38	16.5	353.5	682	9	AJ453748
39	16.5	351.5	724	9	AJ447835
40	16.5	351.5	817	17	CNS03DLV
41	16.4	350.5	668	10	BE005953
42	16.3	349	216	10	AW191803
43	16.3	348.5	531	12	BF426487
44	16.2	347	909	17	CNS03PMF
45	16.2	346.5	1088	13	BM549799

ALIGNMENTS

RESULT 1
AZ820748
LOCUS
DEFINITION
2M0093H12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0093H12 F, DNA sequence.
ACCESSION
AZ820748
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse
Mus musculus

AZ820748 585 bp DNA linear GSS 20-FEB-2001
2M0093H12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0093H12 F, DNA sequence.

AZ820748.1 GI:12990656
GSS.

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 585)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly

M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 CONTACT: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: rdunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0093 row: H column: 12
 Seq primer: CGTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 585.
 Location/Qualifiers
 1..585

FEATURES source

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0093H12"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Ti-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 [G14732114|gb|AF129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 153 a 168 c 117 g 147 t

Alignment Scores:

Pred. No.: 2,98e-84 Length: 585
 Score: 965.00 Matches: 188
 Percent Similarity: 98.45% Conservative: 2
 Best Local Similarity: 97.41% Mismatches: 1
 Query Match: 45.18% Indels: 2
 DB: 17 Gaps: 0

US-09-208-629F-3 (1-407) x AZ820748 (1-585)

Qy 40 AsnValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyPro 59
 Db 11 AATGTTTACACAACTCAGCAACCACTTAACCTATTAAAGAGTTTAAATGGGGTCCC 70
 Qy 60 GluAsnThrPheGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThr 79
 Db 71 CAAATAACCTTTGAAGAATTTCCACATTTCTGACATAGAGGGCTGGACAGGCCACCA 130
 Qy 80 ThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAlaThr 99
 Db 131 ACTATAAACGGAGTGTCGCCAGGACAGATTTTCAACTCTCCACGTGAATAATGTACC 190
 Qy 100 IleGlyThrLeuArgSerLeuSerThrGlnValIleProAlaIleThrIleLeuLeu 119

Db 191 ATAGATACCTCAGAGAAGTTCTTAACTACCAAGTATACCTGCCATCTACATCTCGTGTG 250
 Qy 120 PheVal-ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrly 139
 Db 251 TTTGTGAGTTGGTGATACCAACATCTGTGACCTGTGGAAACTCTCTTAAAGGACCAA 310
 Qy 139 sSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysVa 159
 Db 311 ATCCATCAGTCTGGTTCATCTTTCACCAACTGCCATCGCAGATCTCTTTCTGTGT 370
 Qy 159 IthrLeuProPheLysIleAlaTyHisLeuAsnGlyAsnAsnTrpValPhe-GlyGluV 179
 Db 371 CACACTGCCATTAAAGATCGCTACCATCTCAATGGCAACACTGGGTATTTGGCGGAGG 430
 Qy 179 alMetCysArgIleThrThrValValPheTyGlyAsnMetTyCysAlaIleLeuIleL 199
 Db 431 TCACGTGGCGGATCACCACTGGTGTGTTTCTAGGCAACATGACTGCGCTATCTCTGATCC 490
 Qy 199 eutrCysMetGlyIleAsnArgTyLeuAlaThrAlaHisProPheThrTyGlnLysL 219
 Db 491 TCACCTGCATGGCATCAACCGTACCTGGCCACGGCTCACCTTTCACATACCAAGAGC 550
 Qy 219 euProLysArgSerPheSerLeuLeuMetCysGly 230
 Db 551 TGCCCAAAACGCGCTTCTCCATGCTCATGTGTGGC 585

RESULT 2

AA177828

LOCUS

DEFINITION

5' similar to TR:G663021 G663021 PROTEINASE ACTIVATED RECEPTOR 2.

ACCESSION: AA177828

VERSION: AA177828.1

KEYWORDS: EST.

SOURCE: house mouse.

ORGANISM: Mus musculus

REFERENCE: 1 (bases 1 to 529)

AUTHORS: Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE: The WashU-HMI Mouse EST Project

JOURNAL: Unpublished (1996)

COMMENT: Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:381203

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 494.

Location/Qualifiers

1..529

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:620379"

/clone_lib="Soares mouse 3NBMS"

/sex="male"

/tissue_type="Spleen"

/dev_stage="4 weeks"

/lab_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand

was primed with a Not I - oligo(dT) primer [5,

TGTTACCAATCTGAAGTGGAGCGCGCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
3/], double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTV73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fatima Bonaldo."

BASE COUNT 110 a 168 c 107 g 144 t

Alignment Scores:
Pred. No.: 4,94e-79 Length: 529
Score: 911.00 Matches: 171
Percent Similarity: 99.43% Conservative: 3
Best Local Similarity: 97.71% Mismatches: 1
Query Match: 42.65% Indels: 0
DB: Gaps: 0

US-09-208-629F-3 (1-407) x AAI77828 (1-529)

QY 105 SerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyVal 124
|||||
Db 3 AGTTCTTAAGTACCAAGTGATACCTGCCATCTACATCTCTGCTGTGTGGTTGGTGA 62
|||||
QY 125 ProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuVal 144
|||||
Db 63 CCAGCCAAATCATGTGACCTGTGAAACTCTCTTAAGGACCAATCCATCAGTCGTGTC 122
|||||
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
|||||
Db 123 ATCTTTACACCAACCTGGCCATCTCGAGATCTCTTTCTGTGTGCACATGCCATTTAAG 182
|||||
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyValMetCysArgIleThr 184
|||||
Db 183 ATGCTTACCATCTCAATGGCAACACTGGTATTTGGCGAGGTACGTCGGATCACC 242
|||||
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
|||||
Db 243 ACGGTTGTTTCTACGGCAACATGTACTGGCTATCTGTATCTCTACCTGATGGGCATC 302
|||||
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
|||||
Db 303 AACCGGTACCTGGCCAGGCTCAACCTTTCCACATACAGAGGTGCGCCAAACGACCTTC 362
|||||
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
|||||
Db 363 TCCATGCTCATGTGGCATGGTGTGGTGTATGTTCTTATACATGTCGCCCTTTGTC 422
|||||
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
|||||
Db 423 ATCTGAAGCAGGAGTACCACTCGTCCACTCCGAGATCACCACTGCCACGATGTCGT 482
|||||
QY 265 AspAlaCysGluSerProSerSerPheArgPheTyrTyrPheVal 279
|||||
Db 483 GACGCGTGGCGAGTCCCATCATCTCTCCGATTCTACTACTTCGTC 527
|||||

RESULT 3
BM258063
LOCUS BM258063 581 bp mRNA linear EST 17-DEC-2001
DEFINITION 522739 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BM258063
VERSION BM258063.1 GI:17893662
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 581)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
Perteau,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and

TITLE
JOURNAL
MEDLINE
COMMENT
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGACG
Plate: 126 row: D column: 8
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1..581
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 3BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."
BASE COUNT 122 a 171 c 117 g 171 t
ORIGIN

Alignment Scores:
Pred. No.: 1,99e-66 Length: 581
Score: 782.50 Matches: 140
Percent Similarity: 86.08% Conservative: 27
Best Local Similarity: 72.16% Mismatches: 26
Query Match: 36.63% Indels: 1
DB: Gaps: 1
US-09-208-629F-3 (1-407) x BM258063 (1-581)
QY 100 IleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeu 119
|||||
Db 2 ATGGGGTACTCTGACGAGCCCTTTAAGTACCAATGATACCCGCCATCTACATCTCTGGTG 61
|||||
QY 120 PheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLys 139
|||||
Db 62 TTTGCAGTAGGTATGCGGCCAATGCGGTGACCTGTGGATGCTC---TTCAGGACACGA 118
|||||
QY 140 SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysVal 159
|||||
Db 119 ACCATCCGTATGACCATCTTACACCAACCTGGCCATTCGACAGACTTCTGTTTGTGT 178
|||||
QY 160 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 179
|||||
Db 179 ACATGCGCTTTAGATAGCTTACCATCTCAATGGGAACAACCTGGGTATTGGCGAGGTC 238
|||||
QY 180 MetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeu 199
|||||
Db 239 ATGTGCGGGGCCACACGCTCATCTTATGGCAACATGATTTGTCTCCATTTCTGCTCTC 298
|||||
QY 200 ThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeu 219
|||||
Db 299 GCTGCATCAGTATCAACCGCTACTAGCCATTTGTCATCTCTTCACTACCGGGAGCTG 358
|||||
QY 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr 239
|||||
Db 359 CCCAGCGGACCTACGCTTGTGCATGTGAGCTGTGTGGACACGCTTTCTTTATAC 418
|||||
QY 240 MetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThr 259
|||||

Db 419 ATGCTCGGTTTTTCATCTTGAAGCAGGAGTACTATCTTTGTCAGACACATTAACACC 478
 QY 260 CysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrThrVal 279
 Db 479 TGCACGATGTCACACACATCGAGTCTCTGTCGCCCTTCAGCTCTACTACTTCAATC 538
 QY 280 SerLeuAlaPhePheGlyPheLeuIleProPheValIleIle 293
 Db 539 TCCTTGGCATCTCTTGGATCTTAATCCCATCTTCTGGTCATT 580

RESULT 4
 BM255484 554 bp mRNA linear EST 17-DEC-2001
 LOCUS 517232 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
 DEFINITION
 ACCESSION
 VERSION
 BM255484.1 GI:117891083
 KEYWORDS
 EST.

SOURCE
 ORGANISM
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 554)
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
 G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
 Perteau,G., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and
 Keefe,J.W.

TITLE
 Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)
 21180013
 CONTACT: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390

EMAIL: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCATGACCAT
 BACKWARD: GTTTCGCCAGTCACGACG
 Plate: 117 row: K column: 11
 Seq primer: ATTAGTGACACTATAG.

FEATURES
 source
 1..554
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 3BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."
 BASE COUNT 140 a 148 c 118 g 148 t

ORIGIN
 Alignment Scores:
 Pred. No.: 1-83e-55 Length: 554
 Score: 670.00 Matches: 126
 Percent Similarity: 81.42% Conservative: 23
 Best Local Similarity: 68.85% Mismatches: 32
 Query Match: 31.37% Indels: 2
 DB: 13 Gaps: 2

US-09-208-629F-3 (1-407) x BM255484 (1-554)

QY 15 IleIysMetIysIleIleLeuValAlaAlaGlyLeuLeuPheLeuProValThrVal 34
 Db 9 ATCAAGATGAGCGCGTTCATCTTTGACGCCATTTGGAGCAGTACTTCTGTGCGCTGCTTCC 68

QY 35 CysGlnSerGlyIleAsnVal---SerAspAsnSerAlaLysProThrLeuThrIleLys 53
 Db 69 TGTCAAGCGGCGATCGAATATCATGCACACAACTTGGCAAGCCAACTTGTCTATTAA 128
 QY 54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
 Db 129 ACCTTCCGTGGAGCTCCCAAAATTCCTTTGAAGAGTTCCCCCTTCTGCCCATAGAAGC 188
 QY 74 TTPThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
 Db 189 TGGACAGGAACACCAAAACTGTAATAATCCCTGGAAGAACTTGATTCAATCTC 248
 QY 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
 Db 249 CATGTGAATATGCTACCATGGGTACTCTGACGACCCCTTTAAGTACCAATGTATACC 308
 QY 114 AlaIleTyrIleLeuLeuPheValGlyValProSerAsnIleValThrLeuTrpLys 133
 Db 309 GCATCTACATCTCTGTTGTCAGTAGTATGCCGGCAATGCGTGACCCCTGTGGATG 368
 QY 134 LeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
 Db 369 CTC---TTCAGGACACAGAACCATCGTATGACCATCTTCTACACCAACTGGCCATTGCA 425
 QY 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
 Db 426 GACTTCTGTTTGTGTACACTGCTGCTTTAGATAGTACCATCTCAATGGGAACAC 485
 QY 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
 Db 486 TGGGTATTGGCGAGGTCTATGTCGGCGGCACACCGGTCTCTCTATGGCAACATGTAT 545
 QY 194 CysAlaIle 196
 Db 546 TGCTCCATT 554

RESULT 5
 B0897419

LOCUS
 DEFINITION
 IMAGE:4681483 5', mRNA sequence.

ACCESSION
 B0897419
 VERSION
 B0897419.1 GI:22289433
 KEYWORDS
 EST.
 SOURCE
 African Clawed frog.

ORGANISM
 Xenopus laevis

REFERENCE
 1 (bases 1 to 839)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

AUTHORS
 TITLE
 JOURNAL
 COMMENT

CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Igor Dawid
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
[www-bio.lnl.gov/bbrp/image/image.html](http://www.bio.lnl.gov/bbrp/image/image.html)
 Plate: LLAM10561 row: k column: 20
 High quality sequence stop: 737.

FEATURES
 Location/Qualifiers
 1..839
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:4681483"
 /clone_lib="NICHD XGC Emb4"
 /dev_stage="embryo, stage 31-32"


```

Db      63 GCAGTGTCTGCTTGGAGTAAAGGTCAGAACCGCGCGTGTGTACATGCTGCACCTG 122
Qy      151 AlaileAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsn 170
Db      123 GCATGCGGACGTCTTTCGTGTGGTGTCCCTTCAAGATCAGTACTACTCTCTCC 182
Qy      171 GlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrValValPheTyrGly 190
Db      183 GGCACTGATGGAGTTCGGGTCTGGAATGTGCGCTTGCACCCAGCGTTTACGGG 242
Qy      191 AsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThr 210
Db      243 AACATGTAGCCTCCATCATGCTCATGAGGTATAGCATGACCGTTCTTGGCGGTG 302
Qy      211 AlaHisPro-----PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeu 227
Db      303 GTGTATCCGATCAGTCCCTGTCTGCGGACCTCTGGCGGACCTCAACTTCACT----- 356
Qy      228 MetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLys 247
Db      357 ---TGGTGTCTATTGGGTGATGGCATCATGCGGGTGTGCGCTTCTCTCAAGGAG 413
Qy      248 GlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAlaCys 267
Db      414 CAGACCAACCGAGTTCGGGACTCAACATCACCACCTGCCACGACGCTCTCAGTGAG--- 470
Qy      268 GluSerProSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeu 287
Db      471 AACCTGATGCAAGGCTTTTACTCGTACTACTTCTCGGCGCTTCTCCGCCATCTTCTTT 530
Qy      288 IleProPheValIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSer--- 306
Db      531 GTCCGTTGATGTTCCACGGTCTGCTACACGTCATCATCGCGTCTGCTGCTCTCTCC 590
Qy      307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeu 319
Db      591 GCGGTTGCCAACCGAGCAAGAGTCGGGCTTGT-----TTCGTCTGCGCGG--- 641
Qy      320 LeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHis 339
Db      642 ---GTGTTCTGTCATCTTCATCGTCTGTTGGGCCCCC-AACGTCCTCTGATGTGTCAC 697
Qy      340 HisAlaAsnTyrTyrTyr-----HisAsnThrAspSerLeuTyrPheMetTyr 355
Db      698 -----TACCTTTCTCTCGACAGTCTCGGTACAGAGCAGCTACTTTGTTTAC 748
Qy      356 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheVal 375
Db      749 CTCCTCGCTGTGTGACAGCGTGAGTGTGCTGCTGATCGATCGTTCGTTGATTACTACTAC 808
Qy      376 MetSer 377
Db      809 GCCTCC 814

RESULT 7
CNS04SUP 938 bp DNA linear GSS 24-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 001A08 of library H from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL305674.1 GI:8199451
VERSION AL305674.1
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 938)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bertot,A., Fizames,C., Winkler,P., Brottier,P., Quetier,F.,

```

```

Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 938)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bertot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
3 (bases 1 to 938)
Genoscope.
Direct Submission
Submitted (12-APR-2000)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
Location/Qualifiers
1..938
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="001A08"
/clone_lib="H"
/note="Genoscope sequence ID : XC0AH001BA04XD1-end : T7"
BASE COUNT 161 a 307 c 261 g 207 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 7,45e-35 Length: 938
Score: 462.00 Matches: 89
Percent Similarity: 60.08% Conservative: 57
Best Local Similarity: 36.63% Mismatches: 81
Query Match: 21.63% Indels: 17
DB: 17 Gaps: 4
US-09-208-629F-3 (1-407) x CNS04SUP (1-938)
Qy 145 llePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 23 ATCTTATGCGCAACCTGCGCGTGGC-GACCTGCTCTTCGTCATCTGGGTCCCGTGA 81
Qy 165 lleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 82 ATCGCTTACCCTCAAGGGACACTGGGTCTAGGGAGTCCCTGTGCAAGTCTG 141
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 142 GTGGCTTCTTCTACGGCAACATGTACTGCTCCTCCTTCATCGGTGCATCAGCGTC 201
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 202 CAGCGCTGCAGGCGGTGCTCCACCG-----CTCACAAGAGGGGGG 246
Qy 225 SerLeuLeuMetCysGlyIle-----ValTrpValMetValPheLeuTyrMetLeu 241
Db 247 ACGGTGGCGCGTGGGTGTCCCGGCACTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 306
Qy 242 PropheValIleLeuLysGlnIleTyrHisLeuValHisSerGluIleThrThrCysHis 261
Db 307 CCTGTGTTCTGTGATCAGCAGTCTGTGGTGACAAACCTGGCATCCGACCTGCCAC 366
Qy 262 AspValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu 281
Db 367 GAGTCCAGCCAGCCAGCAGGAGGAGCGGTGCGAGG-----TACTTCTGACCATG 420
Qy 282 AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeuIle 301
Db 421 GGAACTGGGGTATTATGTTCCCGCTCTGTCGTCATCGTCTCTACGCTCTCATGCTC 480
Qy 302 HisLysLeuLysSerLys-----AspArgIleTrpLeuGlyTyrIle 315

```



```

Qy 278 PheValSerLeuAlaPheGlyPheLeuLeuProPheValIleIleIlePheCysTyr 297
Db 501 TTGGTCTGCCTATCGCTGCGCTTCCCTGCTGCTGCGAGTGATGCTTCACTTAC 560
Qy 298 ThrThrLeuIle 301
Db 561 TGCTCAGTGCTA 572

RESULT 12
AL660446
LOCUS AL660446 685 bp mRNA linear EST 13-DEC-2001
DEFINITION AL660446 XGC-neurula silurana tropicalis cDNA clone TNeu043b15 5',
mRNA sequence.
ACCESSION AL660446
VERSION AL660446.1 GI:17674619
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
REFERENCE 1 (bases 1 to 685)
AUTHORS Huckie,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Huckie E
Sanger Centre
Hinckton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu043b15.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
FEATURES
source
location/Qualifiers
1..685
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu043b15"
/clone_lib="XGC-neurula"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/notes="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT 177 a 169 c 152 g 184 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 1.48e-29 Length: 685
Score: 406.00 Matches: 86
Percent Similarity: 57.75% Conservativeness: 37
Best Local Similarity: 40.38% Mismatches: 62
Query Match: 19.01% Indels: 28
DB: 7

US-09-208-629F-3 (1-407) x AL660446 (1-685)

Qy 27 LeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp----- 43
Db 74 GTCCTATTCCTT-----TTGTGTGGGGTGTCTCTGGGGCTCAAGATGAGGATACG 124
Qy 44 -----AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyPro 59
Db 125 CTGAGAGAAATGAATCTCTGCTGCC-----AAGACATTTCCGGGAAGAGAA 172
Qy 60 Gln---AsnThrPheGluGluPheProLeuSerAspIleGluGlyIleThrGlyAlaThr 78
Db 173 GAAGGTGGGGAATATGAGAGCTTTCCCATAGCATCTGCTGCTGCTGCCAGACACCA 232
Qy 79 ThrThr-----IleLysAlaGlu-----CysProGlu 87

```

```

Db 233 ACCAATAATAATCTCTTTTCTAAGAAATCTCCAGCAATAATGCTTTAAAGTCCCCAGCGAA 292
Qy 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
Db 293 CACACCATCACT-----AAAGTCAGCAATTCACCTTGGTGCTACTTAAGTGGCAAGTA 346
Qy 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
Db 347 AGCAGGGAACATGATACCGGGCATCTATATCATTTGTTGTCTCATTTGGTGTGCGCAGCAAT 406
Qy 128 lleValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHis 147
Db 407 GCCTTGGTACTATGATGCTGCTTCATCAGGTGACATCTGTGTCACCATCTGTCTCTAC 466
Qy 148 ThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyr 167
Db 467 GCACGCTTAGCCACCTCTGACTTGTCTATTGCTTCATGCTGCTCTTTAAGATAGCTTAT 526
Qy 168 HisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValVal 187
Db 527 CACTTAAACGGCAACAACACTGGATTTTGGAGAAACCATGTGCGGGCCATGACCATTTTC 586
Qy 188. PheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyr 207
Db 587 TTGTACTTTAACAATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 646
Qy 208 LeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 647 GTTGCCATTGTCCTCATCTTATATACAGGAGCTTGCCG 685

RESULT 13
AJ456719
LOCUS AJ456719 808 bp mRNA linear EST 22-APR-2002
DEFINITION AJ456719 riken1 Gallus gallus cDNA clone 9f2r1, mRNA sequence.
ACCESSION AJ456719
VERSION AJ456719.1 GI:20266815
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 808)
AUTHORS Buerstedde,J.M.
TITLE Gallus gallus bursal lymphocyte EST
JOURNAL Unpublished (2002)
COMMENT Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES
source
location/Qualifiers
1..808
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="9f2r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/notes="CB inbred strain"
BASE COUNT 189 a 169 c 204 g 244 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 3.8e-28 Length: 808
Score: 392.50 Matches: 92
Percent Similarity: 48.81% Conservativeness: 51
Best Local Similarity: 31.40% Mismatches: 97
Query Match: 18.38% Indels: 53
DB: 9

US-09-208-629F-3 (1-407) x AJ456719 (1-808)

```



```

Pred. No.: 1.11e-27 Length: 845
Score: 388.00 Matches: 93
Conservative: 54
Best Local Similarity: 50.87% Mismatches: 92
Query Match: 32.18% Indels: 50
DB: 18.16% Gaps: 8

US-09-208-629F-3 (1-407) x AJ456135 (1-845)

QY 10 AlaGlySerGlnAspIleLeuVal 23 -----1leuLeuLeuVal
Db 55 GCCGGAGCGAGGAGGAGGAGTGGCGGCGCGCGGGCTGCTGCTGCTGCTT 114
QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyLeuValSerAsp 43
Db 115 AGCGCCCTGCTGCGG-----129
QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
Db 130 ---GCTGCCCTCCACAGACAGAGTAGAACACAGAGT-----165
QY 64 GluGluPheProLeuSerAspIleGluGlyThrThrGlyAlaThrThrThrIleLysAla 83
Db 166 -----TCAAAAGGAGAGTGGTCTGGCCAGAGGTTCAGATAGT 210
QY 84 GluCyProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeu 103
Db 211 AACGCCCTCTGAGGAGTACATAA-----GTGATGACTTTCAGCAAAAGTCTT 261
QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
Db 262 ACAGGAAATACTACAGTTCCTTCCACAGTCTATGTCATCTGCTTTATCATGTT 321
QY 124 ValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSer 142
Db 322 TTGCCAAGCATGCTATGCGCATCTGGTCTTTTTCAGACAAAGAGAAATCCT 381
QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 382 GCTGTGATTATATGTTAACTTGGCATTGGCAGACCTTCTTCTGTTGCTGGTCCCA 441
QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 442 CTGAAGATTGCATATCATTTAAATGGAATGCAATGCTGTTGGGAAGGCTCTGCAAA 501
QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMet 202
Db 502 GTGCTTGTGGATTATTTATGGAATATGATGCTGCTCCATCTTTTATGACATGCTC 561
QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
Db 562 AGTGTGCAAGGATTGGGTTGTAGTGCAACCCATAGTGCACTCAAGA---AGGAAGTCT 618
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPr 242
Db 619 GAAATTGCCCTGGGCACTCCCTTGTCTGCTGATATGATGATGTTTGTGGGAACCATCC 678
QY 242 oPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAs 262
Db 679 CCGTATCTCTGCAATCAGACNGTGTATATTTTCAGACCTTAACATCACTACCTGCCATGA 738
QY 262 pValValAlaCysGluSerProSerProSerPhe-----ArgPhe-TyrTyrP 278
Db 739 TGTGTTG-----CCTGAAAATTTTGGCTCATGACATGGTTCAAGTACT 783
QY 278 heValSerLeuAlaPhePheGly 285
Db 784 TCCTCTCACTTTGCCAATTGGGA 806

RESULT 17
AL532537 904 bp mRNA linear EST 13-FEB-2001
LOCUS AL532537 LTI_NFL001_NBC4 Homo sapiens cDNA clone CSODM006YC06 5
DEFINITION
```

```

Db      506 AGGATTGAACATCAAAAGCGGTGCATATTTGCTGGATTCTAGTATTGCTCAGACA 565
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      241 LeuProPheValle-----LeuLysGlnGluTyrHisLeuValHisSerGluille 257
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      566 CTCCTCACTCTCACTCAACCCCTATGTCAAGACGAGG-----GCTGAAAGG 610
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      258 ThrThrCysHisAspValValAlaCysGluSerProSerSerPheArgPheTyrTyr 277
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      611 ATTACATGATGGAGTATCCAACTTTGGAAGAACTAAATCT-----CTTCCTGGATT 664
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      278 PheValSerLeuAlaPhePheGlyPheLeuLeuPheVallellellellePheCysTyr 297
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      665 CTCCTGGGCGCATGTTTCATAGGATATGTTCCACTTATATCATCTTCATCTGCTAT 724
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      298 ThrThrLeulleHisLys-----LeuLysSerLysAsp 308
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      725 TCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACCAAACTCACTGAGAAATCT 784
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      309 ArgileTrpLeuGlyTyrIleLysAlaValLeuLeuLeuVallellellePheThrLysCys 328
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      785 GGTGTAACAAAAGGCTCTCAACACAATATTCTTATTATTGTTGTTGTTCTCTGT 844
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      329 PheAlaProThrAsnIlellelleLeuVallelleHis 339
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      845 TTCACACCTTACCATGTGCAATATTCAACAT 877

```

RESULT 18

```

AL547762      931 bp mRNA linear EST 16-FEB-2001
LOCUS      AL547762 LTI NFL006_PL2 Homo sapiens cDNA clone CS0D1017YN05 5
DEFINITION      prime, mRNA sequence.
ACCESSION      AL547762
VERSION      AL547762.1 GI:12882129
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 931)
AUTHORS      Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 931
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1017YN05"
/clone_lib="LTI NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT)-primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT      268 a 224 c 165 g 273 t 1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      1.96e-27
Score:          386.00
Percent Similarity: 54.18%
Best Local Similarity: 31.87%
Query Match:    18.07%

```

```

DB:      9          Gaps:      5
US-09-208-629F-3 (1-407) x AL547762 (1-931)
Qy      102 TyrLeuArgSerSerLeuSerThrGlnValleProAlaIleTyrIleLeuLeuPheVal 121
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      121 TATGACATCATCAGCAGCGCAGGATAGTAATGCTCTGCATTACAGCCTCGTCTTCATC 180
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerille 141
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      181 ATTGGGCTCGTGGGAACTTACTAGCCTTGCTGCTGCTATTTCAAAACAGGAAAAAATC 240
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      142 ---SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      241 AACTCTACCACCTCTATTCAACAATTTGGTGAATTTCTGATATATCTTTTACCCCGCT 300
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      301 TTGCCTACACGAATAGCCTACTATGCAATGGGCTTTGACTGGAGAAATCGGAGATGCTTG 360
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThr 200
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      361 TGTAGATAAATCGCCTAGTGTTTTACATCAACACATATGCAAGGTGTGAACCTTATGACC 420
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      421 TGCTGAGTATTGACCGCTTCATTCCTGCTGGTGCACCTCTACGCTACACAGATAAAA 480
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      481 AGGATTGAACATGCAAAAGCGGTGCATATTTGCTGCATTCTAGTATTGCTCAGACA 540
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      241 LeuProPheValle-----LeuLysGlnGluTyrHisLeuValHisSerGluille 257
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      541 CTCCTCACTCTCATCAACCTATGTCAAAGCAGGAG-----GCTGAAAGG 585
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      258 ThrThrCysHisAspValValAlaCysGluSerProSerPheArgPheTyrTyr 277
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      586 ATTACATGATGGAGTATCCAACTTTGAGAAACTAAATCT-----CTTCCTGGATT 639
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      278 PheValSerLeuAlaPhePheGlyPheLeuLeuProPheVallellellellePheCysTyr 297
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      640 CTGCTTGGGCGCATGTTTCATAGGATATGTTCTCCACTTATATCATCTCTCATCTAT 699
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      298 ThrThrLeulleHisLys-----LeuLysSerLysAsp 308
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      700 TCTCAGATCTGCTGCAAACTTTCAGAACTGCCAAACCAAACTCACTGAGAAATCT 759
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      309 ArgileTrpLeuGlyTyrIleLysAlaValLeuLeuLeuVallellellePheThrLysCys 328
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      760 GGTGTAACAAAAGGCTCTCAACACAATATTCTTATTATTGTTGTTGTTCTCTCT 819
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy      329 PheAlaProThrAsnIlellelleLeuVallelleHis 339
      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db      820 TTCACACCTTACCATGTGCAATATTATCAACAT 852

```

RESULT 19

```

AJ446632      860 bp mRNA linear EST 19-APR-2002
LOCUS      AJ446632 riken1 Gallus gallus cDNA clone 14d16r1, mRNA sequence.
DEFINITION      AJ446632
ACCESSION      AJ446632
VERSION      AJ446632.1 GI:20213853
KEYWORDS      EST.
SOURCE      Chicken.
ORGANISM      Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE      1 (bases 1 to 860)
AUTHORS      Buerstedde, J.M.
TITLE      Gallus gallus bursal lymphocyte EST
JOURNAL      Unpublished (2002)
COMMENT      Contact: Buerstedde JM

```

Cellular Immunology
Reinrich-Pette-Institute
Martinstr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers

FEATURES
source
1..860
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="14di6r1"
/clone_lib="rikeni"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/note="CB inbred strain"

BASE COUNT 207 a 185 c 204 g 260 t 4 others
ORIGIN

Alignment Scores:
Pred. No.: 2,78e-27 Length: 860
Score: 384.00 Matches: 92
Percent Similarity: 48.68% Conservative: 55
Best Local Similarity: 30.44% Mismatches: 98
Query Match: 17.98% Indels: 57
DB: 9 Gaps: 9

US-09-208-629F-3 (1-407) x AJ446632 (1-860)

QY 10 AlaGlySerGlnAspileLysMetLys-----IleLeuLeuVal 23
Db 18 GCAGGAGCGGGAGGAGGATGCCGGCCGGCTGTGGTCTGCTCTT 77
QY 24 AlaAlaGlyLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 78 AGCGCCCTCTGCCTGCG-----92
QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
Db 93 ---GCTGCCCTCCCACAGACAGACAGTAGTAACACACAGT-----128
QY 64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAla 83
Db 129 -----TCAAAGGAGAAGATTGTGTCGCCAGAGTTCCAGACTAGT 173
QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyLeu 103
Db 174 AACGCCCTCTGAGAGTGATACAAA-----GTGGATGACTTTGCCAGCAAAAGTCCTT 224
QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyIleLeuLeuPheValValGly 123
Db 225 ACAGGAATAAAGTACAGATGTTTTCTCCCATGCTATGTCATGTTTATCATGGT 284
QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
Db 285 TTGCCAAGCAATGCTATGCGCATCTGGTCTTTTTTTCAGAACCAAGAGAACATCCT 344
QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 345 GCTGTGATTATATGTTAACTGTCATGTCGAGACCCTCTCTCGTTGCTGCTCCCA 404
QY 163 PheLysIleAlaTyHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 405 CTGAAGATTGCATATCATTTAAATAGCAATAACTGGCTGTTTGGGGAAGGTCTCTGCAA 464
QY 183 IleThrThrValValPheTyTrpGlyAsnMetTyTrpCysAlaIleLeuThrCysMet 202
Db 465 GTGCTGTGGATTTTTTATGGAATATGACTGCTCATCTTTTATGACATGCTC 524
QY 203 GlyIleAsnArgTyLeuAlaThrAlaHisProPheThrTyGlnLysLeuProLysArg 222
Db 525 AGTGTCGAAGGTATGTTGGTTAGTGAACCCCACTAGTGCACCTAAGA--AGGAAGTCT 581
QY 223 SerPheSerLeuMetCysGlyIleValTrpValMetValPheLeuTyMetLeuPro 242
Db 582 GAAATTCCTGGATCTCCCTGCTATCTGATCTGATCTGATCTGTTGGGAACCATTCG 641

QY 120 PheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLys 139
 Db 601 TTCTCTGGGCTGGCTCCCAACGGCTGGNNCTGGTCTCTGCTACAGGACTGAG 542
 QY 140 SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysVal 159
 Db 541 AAGCTGACCTCCACCATCTCTCTGATGAACCTGGCTGCAGCAGACCTGCTCATCTCA 482
 QY 160 ThrLeuProPheValIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 179
 Db 481 GTCTGCTCCCTCAAGATATCTACTATTCTCTGGGAACAACCTGGCCCTTTGGGAGGT 422
 QY 180 MetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeu 199
 Db 421 CTGTGCGGGTCAACACAGCTTTCTTCTATGGGAACATGCTGCTCAGTGTCTGCTC 362
 QY 200 ThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeu 219
 Db 361 ACCTGTCATGATGTGGACGATACCTGGGGTGTGTCATCTCTCTCTGGGCTCTTTC 302
 QY 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr 239
 Db 301 CGTACCCCACTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
 QY 240 MetLeuProPheValIleLeuLysGlnIleTyrHisLeuValHisSerGluIleThrThr 259
 Db 241 ACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 182
 QY 260 CysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyr----- 277
 Db 181 TGCCACGATGTTCTCCCTGCTTACACACTGCCAGGCTGCTTTTGTAGTACACCCAGGT 122
 QY 278 -----PheValSerLeuAlaPhePheGlyPheLeuLeuPro 289
 Db 121 CTGCTTTTATGATGCTGTGAGCTGCTGCTATCTGCTATCTGCTTCTATTCATTT----- 71
 QY 290 PheValIleIleIlePheCysTyr 297
 Db 70 -----TTCTGCTAC 62
 RESULT 21
 AJ456491 760 bp mRNA linear EST 22-APR-2002
 LOCUS AJ456491 riken1 Gallus gallus cdna clone 816r1, mRNA sequence.
 DEFINITION AJ456491
 ACCESSION AJ456491
 VERSION AJ456491.1 GI:20266587
 KEYWORDS EST.
 SOURCE chicken.
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 760)
 Buerstedde,J.M.
 Gallus gallus bursal lymphocyte EST
 Unpublished (2002)
 Contact: Buerstedde JM
 Cellular Immunology
 Heinrich-Pette-Institute
 Martinistr. 52, 20251 Hamburg, Germany
 Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.
 Location/Qualifiers
 1..760
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /clone="816r1"
 /clone_lib="riken1"
 /cell_type="bursal lymphocyte"
 /dev_stage="2-3 weeks old"
 /notes="CB inbred strain"
 179 a 165 c 197 g 217 t 2 others

FEATURES

source

RESULT 22

AJ451396

LOCUS

DEFINITION

AJ451396

ACCESSION

AJ451396.1

VERSION

KEYWORDS

EST.

Alignment Scores:

Pred. No.: 3,336-27 Length: 760
 Score: 382.50 Matches: 83
 Percent Similarity: 51.53% Conservative: 52
 Best Local Similarity: 31.68% Mismatches: 88
 Query Match: 17.91% Indels: 39
 DB: 9 Gaps: 6

US-09-208-629F-3 (1-407) x AJ456491 (1-760)

QY 10 AlaGlySerGlnAspIleLysMetLys-----IleLeuLeuVal 23
 Db 63 GCGGAGCGGAGGAGGAGGAGGATGCCGCGCGCGCGCTGTGCTGTGCTGTGCTGT 122
 QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
 Db 123 ACGGCCCTGCTGGCG----- 137
 QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
 Db 138 ---GCTGCGCTCCACAGACAGACAGTAGAACACGAGT----- 173
 QY 64 GluGluPheProLeuSerAspIleGluGlyTyrThrGlyAlaThrThrThrIleLysAla 83
 Db 174 -----TCAAAAGGAAGAAGTTTGTGTCGCCAGAGGTTCCAGATACAGT 218
 QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
 Db 219 AAGCGCTCTGAGAGTGTACACAAA-----GTGGATGACTTTGCGACCAAGTCCIT 269
 QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGly 123
 Db 270 ACAGAGAAACAACTACAGTTTCTCCCACTGCTATGTCATGTCATTATCATTTGT 329
 QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
 Db 330 TTGCCAAGCAATGCTATGCCCATCTGGGTCTTTTTTTCAGAACAAAGAAACATCCT 389
 QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
 Db 390 GCTGTGATTATGCTTAACTTGGCATGGCAGACCTTCTCTTGTGTTGTTCCCA 449
 QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
 Db 450 CTGAAGATTGCATATCATTTAAATGCAATAACTGCTGTTGGGGAAGGTCTCTGCAAA 509
 QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMet 202
 Db 510 GTGCTTTGTTGGATTTTATGAAAATATGTAAGTCTGCTCCATTTCTTTATGACATGCTC 569
 QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
 Db 570 AATGTGCAAGGATTTGGTTGTAGTGAACCCCATAGTCGACTCAAGA---AGGAAGTCT 626
 QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
 Db 627 GAAATGCTGCTGGCATCTCCCTTCTGATCTGATCTGATTTGTTGGGAACCATTCG 686
 QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
 Db 687 CTGTATCTTGTCAATCAGACNGGNGTATTTTAGACCTTTAATCATCACTACCTGCTCATGAT 746
 QY 263 ValVal 264
 Db 747 GTGTGT 752

RESULT 22
 AJ451396
 LOCUS
 DEFINITION
 AJ451396
 ACCESSION
 AJ451396.1
 VERSION
 KEYWORDS
 EST.

773 bp mRNA linear EST 22-APR-2002
 LOCUS AJ451396 riken1 Gallus gallus cdna clone 2816r1, mRNA sequence.

DEFINITION
 AJ451396
 ACCESSION
 AJ451396.1
 VERSION
 KEYWORDS
 EST.


```

SOURCE      chicken.
ORGANISM    Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 773)
AUTHORS     Buerstedde,J.M.
TITLE       Gallus gallus bursal lymphocyte EST
JOURNAL     Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinstr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES
  source
    1..773
    /organism="Gallus gallus"
    /db_xref="taxon:9031"
    /clone_lib="rikeni"
    /cell_type="bursal lymphocyte"
    /dev_stage="2-3 weeks old"
    /note="CB inbred strain"
  BASE COUNT  182 a  165 c  190 g  235 t
  ORIGIN
    Alignment Scores:
    Pred. No.:      3,41e-27      Length:      773
    Score:          382.50        Matches:     86
    Percent Similarity: 51.88%    Conservative: 52
    Best Local Similarity: 32.33% Mismatches:    82
    Query Match:    17.91%       Indels:       46
    DB:             9            Gaps:         8
  US-09-208-629F-3 (1-407) x AJ451396 (1-773)
  QY  10 AlaGlySerGlnAspLeuLysMetLys-----lleLeuLeuLeuVal 23
  Db  19 GCCGGAGGAGGAGGAGGAGGAGGATGCCGGCGCGCGGCTGTGCTGTGCTGCTT 78
  QY  24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyLeuValSerAsp 43
  Db  79 AGCGCCCTGTGGCG-----
  QY  44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
  Db  94 ---GCTGCCCTCCACAGACACAGTAGAACACAGCAGT-----
  QY  64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAla 83
  Db  130 -----TCAAAAGGAAGAAGTTTGTGTCGCGCAGAGGTTCCAGATAGT 174
  QY  84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTrpLeu 103
  Db  175 AACGCCCTCTCAGAGTCATACAAA-----GTGGATGACTTTCAGCAAAAGTCCTT 225
  QY  104 ArgSerSerLeuSerThrGlnValIleProAlaIleTrpIleLeuLeuPheValValGly 123
  Db  226 ACAGGAAACTACTACAGTTTTTCTCCACAGTGTCTATGTCATTTATCATGGT 285
  QY  124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
  Db  286 TTGCCAAGCAATGCTATGGCCATCTGGTCTTTTTTTTCAGAAAGAAAGAACATCCT 345
  QY  143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
  Db  346 GCTGTGATTATATGTTAATCTGGCATTTGGCAATTAAGTGGCTGTTTGGGAGAGTCTCTG 405
  QY  163 PheLysIleAlaTrpHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
  Db  406 CTGAGATTTCATATCATTTAATGGCAATTAAGTGGCTGTTTGGGAGAGTCTCTGCAAA 465
  QY  183 IleThrThrValValPheTrpGlyAsnMetTrpCysAlaIleLeuLeuThrCysMet 202

```

```

  Db  466 GTGCTGTGTGGATTTTATGGAATATGTACTGCTCCATTTCTTTTATGACATGCTC 525
  QY  203 GlyIleAsnArgTrpLeuAlaThrAlaHisProPheThrTrpGlnLysLeuProLysArg 222
  Db  526 AGTGTGCAAAAGTATGGGTTGTAGTGAACCCCATAGTCACATCA-----AGAAGG 576
  QY  223 SerPheSerLeuLeuMetCysGlyIle-----Val-TrpValMetValPheLeuTrp 239
  Db  577 AAGTTTGAATTTGCCCTG---GCCATCTCCCTTGTGATCTTGGATCTGATTTGTGGG 633
  QY  239 rMetLeuProPheValIleLeuLysGlnGluTrpHisLeuValHisSerGluIleThrTh 259
  Db  634 AACCATTCGCGTGTATCTTGTCAATCAGACNGTGTATATTTCAGACCTTTAAACATCACTAC 693
  QY  259 rCysHisAspValVal 264
  Db  694 CTGCCATGATGTGTG 709
  RESULT 23
  AJ456561
  LOCUS      890 bp mRNA linear EST 22-APR-2002
  DEFINITION riken1 Gallus gallus cDNA clone 807r1, mRNA sequence.
  ACCESSION AJ456561
  VERSION   AJ456561.1 GI:20266657
  KEYWORDS  EST.
  SOURCE     chicken.
  ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
  REFERENCE 1 (bases 1 to 890)
  AUTHORS    Buerstedde,J.M.
  TITLE      Gallus gallus bursal lymphocyte EST
  JOURNAL    Unpublished (2002)
  COMMENT    Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinstr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
  FEATURES
    source
      1..890
      /organism="Gallus gallus"
      /db_xref="taxon:9031"
      /clone_lib="rikeni"
      /cell_type="bursal lymphocyte"
      /dev_stage="2-3 weeks old"
      /note="CB inbred strain"
  BASE COUNT  211 a  185 c  219 g  270 t
  ORIGIN
  Alignment Scores:
  Pred. No.:      4,07e-27      Length:      890
  Score:          382.50        Matches:     84
  Percent Similarity: 51.71%    Conservative: 52
  Best Local Similarity: 31.94% Mismatches:    87
  Query Match:    17.91%       Indels:       40
  DB:             9            Gaps:         6
  US-09-208-629F-3 (1-407) x AJ456561 (1-890)
  QY  10 AlaGlySerGlnAspIleLysMetLys-----lleLeuLeuLeuVal 23
  Db  55 GCCGGAGCGGAGGAGGAGGAGGATGCCGGCGCGCGGCTGTGCGTGTGCTGCTT 114
  QY  24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
  Db  115 AGCGCCCTCTGTGGCG-----
  QY  44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
  Db  130 ---GCTGCCCTCCACAGACACAGTAGAACACAGCAGT-----

```

```

QY 64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrThrIleLysAla 83
Db 166 -----TCAAAGGAAGAGTTTGTGGCCAGAGGTTCCAGATACGT 210

QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeu 103
Db 211 AACGCCTCTCAGGAGTCATACAAA-----GTGGATGACTTTCAGCAAAAGTCCTT 261

QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
Db 262 ACAGGAAACAACTACAGTTTTCTTCCACAGTCTATGTCATCTCTTTATCATTTGTT 321

QY 124 ValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSer 142
Db 322 TTGCCAAGCANTGCTATGGCCATCTGGGCTTTTTTTCAGAAAGAGAAAGAACATCCT 381

QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 382 GCTGTGATTATATGGTTAACTTGGCATTGGCAGACCTTCTCTCTGTTGTTGCCA 441

QY 163 PheIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 442 CTGAAGATTGCATCATATTTAAATGGCAATAAAGTGGTGTGGGAAAGTCTCTCCAAA 501

QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMet 202
Db 502 GTGCTGTGGATTATTTATGAATAATGATGATCTCTCCATCTTTTATGACATGTC 561

QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
Db 562 AGTGTGCAAGGATTATGGTGTAGTGAACCCCATAGTGCACCTCAAGA---AGGAAGTCT 618

QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 619 GAAATTCGCCCTGGGCATCTCCCTTGCTATCTGGATACATGATTTTGTGGAAACCATCCG 678

QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSer---GluIleThrThrCysHisAs 262
Db 679 CTGTATCTGTCAATCAGACGGTGTATATTTTCAGACCTTAACATCACTACCTGCCATGA 738

QY 262 pValVal 264
Db 739 TGTGTTG 745

RESULT 24
AJ446611 747 bp mRNA linear EST 19-APR-2002
LOCUS riken1 Gallus gallus cdna clone 14c16r1, mRNA sequence.
ACCESSION AJ446611
VERSION AJ446611.1 GI:20213832
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 747)
Buerstedde, J.M.
Gallus gallus bursal lymphocyte EST
Unpublished (2002)
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers
1..747
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="14c16r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
FEATURES
source

```

```

/dev_stage="2-3 weeks old"
/note="CB inbred strain"
BASE COUNT 178 a 158 c 187 g 224 t
ORIGIN
Alignment Scores:
Pred. NO.: 4.09e-27 Length: 747
Score: 381.50 Matches: 82
Percent Similarity: 51.35% Conservative: 51
Best Local Similarity: 31.68% Mismatches: 87
Query Match: 17.86% Indels: 39
DB: 9 Gaps: 6
US-09-208-629F-3 (1-407) x AJ446611 (1-747)
QY 10 AlaGlySerGlnAspIleLysMetLys-----IleLeuLeuVal 23
Db 18 GCCGAGCGGAGGAGGAGGATGCCGCGCGCGCGCTGTGCGTGTGCTGTGCTGCT 77
QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 78 ACGGCCCTGCTGGCG-----
QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
Db 93 ---GCTGCGCTCCACAGACAGACAAGTAGAACACAGCAGT-----128
QY 64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrIleLysAla 83
Db 129 -----TCAAAGGAAGAGTTTGTGGCCAGAGGTTCCAGATACGT 173
QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeu 103
Db 174 AACGCCTCTCAGGAGTCATACAAA-----GTGGATGACTTTCAGCAAAAGTCCTT 224
QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
Db 225 ACAGGAAACAACTACAGTTTTCTTCCACAGTCTATGTCATCTCTTTATCATTTGTT 284
QY 124 ValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSer 142
Db 285 TTGCCAAGCANTGCTATGGCCATCTGGGCTTTTTTTCAGAAAGAGAAAGAACATCCT 344
QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 345 GCTGTGATTATATGGTTAACTTGGCATTGGCAGACCTTCTCTCTGTTGTTGCCA 404
QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 405 CTGAAGATTGCATCATATTTAAATGGCAATAAAGTGGTGTGGGAAAGTCTCTGCCAAA 464
QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMet 202
Db 465 GTGCTGTGGATTATTTTATGGAATAATGATGATCTCTCCATCTTTTATGACATGCTC 524
QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
Db 525 AGTGTGCAAGGATTATGGGTTAGTGAACCCCATAGTGCACCTCAAGA---AGGAAGTCT 581
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 582 GAAATTCGCCCTGGGCATCTCCCTTGTATCTGATACACTGATTTTGTGGAAACCATTCG 641
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHis 261
Db 642 CTGTATCTGTCAATCAGACGGTGTATATTTTCAGACCTTTAAACATCACTACCTGCCAT 698
RESULT 25
AL551903 946 bp mRNA linear EST 16-FEB-2001
LOCUS AL551903 LTI_NFL006.PL2 Homo sapiens cdna clone CS0D1060Y018.5
DEFINITION prime, mRNA sequence.
ACCESSION AL551903

```

```

VERSION      AL551903.1  GI:12890291
KEYWORDS     EST.
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 946)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES     Location/Qualifiers
             1..946
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="LTI_NFL006_PL2"
                /clone="CS0DI060YJ18"
                /tissue_type="placenta"
                /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT   269 a 231 c 168 g 275 t 3 others
ORIGIN
Alignment Scores:
Pred. No.:      6,16e-27      Length:      946
Score:          381.00      Matches:      80
Percent Similarity: 53.39%      Conservative: 54
Best Local Similarity: 31.87%      Mismatches: 97
Query Match:    17.84%      Indels:      20
DB:              9          Gaps:        5

US-09-208-629f-3 (1-407) x AL551903 (1-946)

Qy 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheVal 121
Db 127 TATGCACATCACAGCAGCGCAGATAGTATGCTCTGCATTACAGCTCGCTTCATC 186
Qy 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerIle 141
Db 187 ATTGGGCTCGTGGGAACTTACTAGCTTGCTGCTGCTTCAAAACAGGAAAAATC 246
Qy 142 ---SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThr 160
Db 247 AACTCTACCACTCTCTATTCAACAAATTTGGTGATTTCTGATATACTTTTACCAACCGCT 306
Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 307 TTGCTTACAGTATAGCTTACTATGCAATGGGCTTTGACGTGGAGATCGGAGATCGCTTG 366
Qy 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
Db 367 TGTAGGATAACTGCGVAGTGTTTTACATCAACATATGCAGGTGTGAACCTTATGACC 426
Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 427 TGCTGTAGTATGACCGCTTCAATGTGTGGTGGTGCACCTCTACGCTACACAGATAAAA 486
Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 487 AGGATTGAACATGCAAAAGCGGTGCATATTTGCTGGATTTCTAGTATTGCTGCAGACA 546
Qy 241 LeuProPheValIle-----LeuLysGlnGluTyrHisLeuValHisSerGluIle 257

```

```

Db 547 CTCCCACTCTCTCATCAACCTATGTCAAGCAGGAG-----CTGAAAGG 591
Qy 258 ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 592 ATTACATGATGATGAGGATATCCAAACTTTGAGAGAACTAAATCT-----CTTCCTGGATT 645
Qy 278 PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
Db 646 CTGCTTGGGGCATGTTTCATAGGATATGTACTTCCACTTATAATCATCTCATCTGCTAT 705
Qy 298 ThrThrLeuIleHisIys-----LeuLysSerLysAsp 308
Db 706 TCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAACCCACTCACTGAGAAATCT 765
Qy 309 ArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCys 328
Db 766 GGTGTAAACAAAAGGCTCTCAACACAATATTCTTATTATTGTTGTTGTTGTTCTCTCT 825
Qy 329 PheAlaProThrAsnIleIleLeuValIleHis 339
Db 826 TTCACACCTTACCATGTTGCAWTTATTCAACAT 858

RESULT 26
BC013202      2020 bp      mRNA      linear      HTC 29-AUG-2001
LOCUS         Homo sapiens, clone IMAGE:4214482, mRNA.
DEFINITION    BC013202
ACCESSION     BC013202.1  GI:15342000
VERSION       HTC.
KEYWORDS      Homo sapiens.
SOURCE        Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 2020)
AUTHORS       Strausberg, R.
TITLE         Direct Submission
JOURNAL       Submitted (27-AUG-2001) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK        NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT       Contact: MGC help desk
              Email: cgapbs@mail.nih.gov
              Tissue Procurement: David N. Louis, M.D.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Baylor College of Medicine Human Genome
              Sequencing Center
              Center code: BCM-HGSC
              Web site: http://www.hgsc.bcm.tmc.edu/cdna/
              Contact: villalob@bcm.tmc.edu
              Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
              A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
              Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Place: 26 Row: h Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 2695873
This clone has the following problem: frame shifted.
FEATURES      Location/Qualifiers
             1..2020
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4214482"
                /tissue_type="Brain, anaplastic oligodendroglioma with
1p/19q loss"
                /clone_lib="NCI CGAP_Brn67"
                /lab_host="DH10B"
                /note="Vector: pCMV-SPORT6"
BASE COUNT   410 a 689 c 508 g 413 t

```

ORIGIN

Alignment Scores:

Pred. No.:	3,538-26	Length:	2020
Score:	377.50	Matches:	99
Percent Similarity:	45.43%	Conservative:	60
Best Local Similarity:	28.29%	Mismatches:	136
Query Match:	17.67%	Indels:	56
DB:	11	Gaps:	9

US-09-208-629F-3 (1-407) x BC013202 (1-2020)

Qy	42	SerAspAsnSerAlaIysProThrLeuThrIleIysSerPheAsnGly-----Gly	58
Db	86	TCTGACTCCAGC-----CAAAGCATGAATGCCTTGAAGTGGCT	124
Qy	59	ProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGlyTyrThrGlyAlaThr	78
Db	125	CCCCAGGTCTGATCACCMACTTCTCCCTGGCCAGCGCAGAG-----	166
Qy	79	ThrThrIleIysAlaGlyCysProGluAspSerIleSerThrLeuHisValAsnAsnAla	98
Db	167	-----CAATGTGCCAGGAG-----	181
Qy	99	ThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeu	118
Db	182	-----ACGCCATCGAGAACATGCTGTTCGGCTCCTTCTACCTTCTG	223
Qy	119	LeuPheValValGlyValProSerAsnIleValThrLeuTyrPlysLeuSerLeuArgThr	138
Db	224	GATTTTATCCTGGCTTTAGTTGGCAATACCTCGTCTGTGGCTTTTCATCCGAGACCA	283
Qy	139	LysSer---IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe	157
Db	284	AAGTCGGGACCCCGCCACGCTGTTCTGTGATGCATCTGGCGTGGCGGCGACTTGT	343
Qy	158	CysValThrLeuProPhePylsIleAlaTyrHisLeuAsnGlyAsnAsnTyrPvalPheGly	177
Db	344	GTGCTGGTCTGCCACCGCGCTGTCTACCACTTCTCTGGAACCACTGCGCAATTGGG	403
Qy	178	GluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu	197
Db	404	GAATCGCATGGCGTCTACCGGCTTCCTCTTCTACCTCAACATGTAGCCAGCATCTAC	463
Qy	198	IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln	217
Db	464	TTCTCTCACTGTGATCAGCGCCGACCGTTCCTGGCCATTGTGCACCCGGTCAAGTCCCTC	523
Qy	218	LysLeuProIysArgSerPheSerLeuLeuMetCysGlyIleValTyrPvalMetValPhe	237
Db	524	AAGCTCCGAGGCCCTCTACGCACACCTGGCTGTGCTTCTCTGTGGTGGTGGTGGCT	583
Qy	238	LeuTyrMetLeuProPheValIleLeuIysGlnGluTyrHisLeuValHisSerGluIle	257
Db	584	GTGGC-ATGGCCCGCTGCTGTGTGAGCCACAG-----ACCGTG	621
Qy	258	ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr	277
Db	622	CACAGCAACCAACGCTGTGCTGCTGCAGCTGTACCGGGAGAGGGCTCCCCACCATGCC	681
Qy	278	PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr	297
Db	682	CTGGTGTCCCTGGCA---GTGGCCCTTACCTTCCCGTTTCATCACCAGGTCACCTGCTAC	738
Qy	298	ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTyrPLeuGly	313
Db	739	CTGCTGATCATCCGACGCTCGCGGAGGGCGCTGCTGTGGAGAGCGCCTCAAGACCAAG	798
Qy	314	TyrIleIysAlaValLeuLeuIleValIlePheThrIleCysPheAlaProThrAsn	333
Db	799	GCAGTCGGCATGATGCCCATGTGTGGCCATTTCTTCGTCTCTCTGTGGCTTACCAC	858
Qy	334	IleIleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSfs-----	350

QY 345 TyHisAsnThrAspSerLeuTyPheMetTyLeuIleAlaLeuCysLeuGlySerLeu 364
 Db 1315 TGGCACACCTCAACGCCATCAACATGGGATATAGATACCCGGCGGTGGCCAGCGCC 1374
 QY 365 AsnSerCysLeuAspProPheLeuTyPheValMetSerLysValVal 380
 Db 1375 AACAGTGTCTTACCCCGGTACTCTACTCTTCTGGCAGGCGAGAGACTTGTCCGCTTTGCC 1434
 QY 381 ---AspGlnLeuAsnPro***SerAlaMetAlaArgPro----- 392
 Db 1435 CGAGATGCCAAGCCACCGAGCGAGCTACCCCGAGCCACAGGCTCGTCGCAAGTGGGC 1494
 QY 393 Leu**ArgProArgAspIleTyrGluAspIle 404
 Db 1495 CTCACAGGCTAACAGAACTGTGAGGAAGATTG 1530
 RESULT 28
 LOCUS AK005013
 DEFINITION Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone1300015C04:purinergic receptor p2Y, G-protein coupled 2, full insert sequence.
 ACCESSION AK005013.1 GI:12836638
 VERSION
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, H., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genom. Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20493174
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Ito, Y., Iizawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genom. Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,

Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringuet, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851
 REFERENCE 5 (bases 1 to 3001)
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Iizawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
 COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGCGCGCCGACCTCGATCTTTTITTTTITTTN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGATCCAGAGCTCAATTAATTAACCCGCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.
 FEATURES
 Location/Qualifiers
 1..3001
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:1300015C04"
 /db_xref="MGD:MG1:1895057"
 /db_xref="taxon:10090"
 /clone="1300015C04"
 /sex="male"
 /tissue_type="liver"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 1..3001
 /gene="P2ry2"
 477..1598
 /gene="P2ry2"
 /note="data source:MGD, source key:MG1:105107, evidence:ISS
 purinergic receptor p2Y, G-protein coupled 2 putative"
 /codon_start=1
 /protein_id="BAB23746.1"
 /db_xref="GI:12836639"
 /db_xref="MGD:MG1:105107"
 /translation="MAADLEPNWSTINGTWEGDELGYKCRFNEPKYLLPVSYGVVC
 VLGLCLNVVALYIFLCRLKTNASTYFHLAVSDLSYAASLPLLYVYYARGDHPFS

Email: URL: <http://genetics.hpi.uni-hamburg.de/dt4dest.html>.

FEATURES

Location/Qualifiers
 1. 740
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /clone="27a5r1"
 /clone_lib="rikeni"
 /cell_type="bursal lymphocyte"
 /dev_stage="2-3 weeks old"
 /note="CB inbred strain"
 BASE COUNT 176 a 158 c 184 g 219 t 3 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1,95e-26 Length: 740
 Score: 374.50 Matches: 83
 Percent Similarity: 51.33% Conservative: 52
 Best Local Similarity: 31.56% Mismatches: 88
 Query Match: 17.53% Indels: 40
 DB: 9 Gaps: 6

US-09-208-629F-3 (1-407) x AJ450921 (1-740)

QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
 DB 69 AGCGCCTGCTGGC----- 83
 QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
 DB 84 --GCTGCCCTCCACAGACACAAGTAGAACACGAGT----- 119
 QY 64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrIleLysAla 83
 DB 120 -----TCAAAAGGAAGAAGTTTCTGCGCAGAAAGTTCCAGATACAGT 164
 QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
 DB 165 AAGCCTCTGAGAGTCATCAAA-----GTGATGACCTTCAGCAAAAGTCCTT 215
 QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
 DB 216 ACAGGAAACCTAACTACAGTTTTCTTCCCACTGCTATGTCATTGCTTTATCATTTGGT 275
 QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
 DB 276 TTGCCAAGCAATGCTATGGCCATCTGGTCTTTTTTTCAGAAACAAGAACATCCT 335
 QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
 DB 336 GCTGTGATTATATGTTAACTGGCATTTGGCAGACTCTCTGCTGCTGGTCCCA 395
 QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
 DB 396 CTGAAGATTGCATATCATTTAAATGGCAATACTGGCTGTTGGGAAGTCTCTGCAAA 455
 QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMet 202
 DB 456 GTGCTTATTGG-ATTTTATGGAAATATGATGCTCCATCTTTTATGACATGCTC 514
 QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
 DB 515 AGTGTGCAAGGATTGGTTGGTTAGTGAAACCCATAGTCACCAAGA---AGGAAGTCT 571
 QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
 DB 572 GAAATTGCCCTGGCATCTCCCTGCTATCTGATGACTGATTGTTGGGAACCATCCG 631
 QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
 DB 632 CTGTATCTTCTCAATCAGACNGGGTATATTTCAAACCTTAACATCATCTACCTGCCATGAT 691
 QY 263 Val 263
 DB 692 GTG 694

RESULT 30

LOCUS AJ450921 740 bp mRNA linear EST 19-APR-2002
 DEFINITION AJ450921 riken1 Gallus gallus cdna clone 27a5r1, mRNA sequence.
 ACCESSION AJ450921
 VERSION AJ450921.1 GI:20218142
 KEYWORDS EST.
 SOURCE chicken.

ORGANISM

Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 740)

REFERENCE

AUTHORS Buerstedde, J.M.
 TITLE Gallus gallus bursal lymphocyte EST
 JOURNAL Unpublished (2002)
 COMMENT Contact: Buerstedde JM
 Cellular Immunology
 Heinrich-Pette-Institute
 Martinistr. 52, 20251 Hamburg, Germany

US-09-208-629F-3 (1-407) x AJ450835 (1-766)

```
QY 10 AlaGlySerGlnAspIleLysMetLys-----IleLeuIleLeuVal 23
Db 40 GCCGGAGCGAGGAGGAGGATGGCGGCGCGCGGCTGTGCTGTCTGTCTGCTT 99
QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 100 AGCGCCCTGTCTGGCG----- 114
QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyClyProGlnAsnThrPhe 63
Db 115 ---GCTGCCCTCCACAGACAGTAGAACACGAGT----- 150
QY 64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrIleLysAla 83
Db 151 -----TCAAAAGGAGAGTGTTCGCGCAGAGGTTCCAGATACGTAGT 195
QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
Db 196 AACGCTCTCAGGAGTACATAA-----GTGATGACTTTGCAGCAAAAGTCCTT 246
QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
Db 247 ACAGGAAACATAACTACAGTCTTCTCCACAGTCTGTATGTCATCTCTTTATCATTTG 306
QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
Db 307 TTGCCAAGCAATGCTATGGCCATCTGGGCTCTTTTTCAGAACAAAGAAACATCCT 366
QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 723 GTGTG 728
QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 427 CTGAAGATTTCATATCATTTAAATGGCAATAACTGCTGTTCGGAAGTCTCTGCAAA 486
QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMet 202
Db 487 GTGCTGTGGATTTTATGAATAATATGATCTCTCCATCTCTTTTATGACATGCTC 546
QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
Db 547 AGTGTGCAAAAGTATGGGTGTAGTGAAACCCATAGTGCACTCAAGA---AGGAAGTCT 603
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 604 GAAATTGCCCTGGGCATCTCCCTTGCTATCTGGATACGTATTTTGTGGAAACCATTCG 663
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
Db 664 CTGTATCTTGTCAATCAGACNGTGATATATTC-AGACCTTACATCATCTCTGCCATGAT 722
QY 263 ValVal 264
Db 723 GTGTG 728
```

RESULT 33
AJ450259 746 bp mRNA linear EST 19-APR-2002
LOCUS
DEFINITION AJ450259 riken1 Gallus gallus cDNA clone 24k5r1, mRNA sequence.
ACCESSION
VERSION AJ450259.1 GI:20217480
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Buerstedde, J.M. Phasianinae; Gallus.
Gallus gallus bursal lymphocyte EST

JOURNAL
COMMENT

Unpublished (2002)
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.

FEATURES
source

1. 746
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="24k5r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/note="CB inbred strain"
176 a 155 c 191 g 220 t 4 others

BASE COUNT
ORIGIN

Alignment Scores:
Pred. No.: 1.19e-25 Length: 746
Score: 366.50 Matches: 82
Percent Similarity: 51.74% Conservative: 52
Best Local Similarity: 31.66% Mismatches: 85
Query Match: 17.16% Indels: 40
DB: 9 Gaps: 6

US-09-208-629F-3 (1-407) x AJ450259 (1-746)

```
QY 10 AlaGlySerGlnAspIleLysMetLys-----IleLeuIleLeuVal 23
Db 40 GCCGGAGCGAGGAGGAGGATGGCGGCGCGGCTGTGCTGTCTGTCTGCTT 99
QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 100 AGCGCCCTGTCTGGCG----- 114
QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyClyProGlnAsnThrPhe 63
Db 115 ---GCTGCCCTCCACAGACAGTAGAACACGAGT----- 150
QY 64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrIleLysAla 83
Db 151 -----TCAAAAGGAGAGTGTTCGCGCAGAGGTTCCAGATACGTAGT 195
QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
Db 196 AACGCTCTCAGGAGTACATAA-----GTGATGACTTTGCAGCAAAAGTCCTT 246
QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
Db 247 ACAGGAAACATAACTACAGTCTTCTCCACAGTCTGTATGTCATCTCTTTATCATTTG 306
QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
Db 307 TTGCCAAGCAATGCTATGGCCATCTGGGCTCTTTTTCAGAACAAAGAAACATCCT 366
QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 367 GCTGTGATTTATGATGTTAACTTGGCATTTGGCAGACCTTCTCTGCTGTCTGTTCCCA 426
QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 427 CTGAAGATTGCATATCATTTAAATGGCAATAACTGCTGTTCGGAAGTCTCTGCAAA 486
QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMet 202
Db 487 GTGCTGTGGATTTTATGAATAATATGATCTCTCCATCTCTTTTATGACATGCTC 546
QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArg 222
Db 547 AGTGTGCAAAAGTATGGGTGTAGTGAAACCCATAGTGCACTCAAGA---AAGTCT 603
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
```


Berlin-Charlottenburg, GERMANY; Email: clone@zpd.de.

FEATURES

source

Location/Qualifiers
1. .495
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp686N1799"
/clone_lib="686 (synonym: hlcc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/notes="vector: pRip1Ex2; Site_1: SfIIA; Site_2: SfiIB;
cDNA-collection"

BASE COUNT 132 a 128 c 107 g 128 t
ORIGIN

Alignment Scores:

Pred. No.: 2,74e-25 Length: 495
Score: 360.50 Matches: 72
Percent Similarity: 78.90% Conservative: 14
Best Local Similarity: 66.06% Mismatches: 22
Query Match: 16.88% Indels: 1
DB: 9 Gaps: 1

US-09-208-629F-3 (1-407) x AL713459 (1-495)

Qy 11 GlySerGlnAspIleLysMetLysIleLeuIleLeuValAlaAlaGlyLeuLeuPheLeu 30
Db 168 GGACTCAGTCAATCAAAATGAAGCCCTCATCTTTCAGCTCTGCTCCCTCTTCTG 227
Qy 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlaLysProThr 49
Db 228 TTGCCACTTTTGTGTCAGAGTGCATGGAATATGATACAAACAACTTGGCAAGCCAAAC 287
Qy 50 LeuThrIleLysSerPheAsnGlyProGlnAsnThrPheGluGluPheProLeuSer 69
Db 288 TTACCATTAAGACCTTTCTGTGAGCTCCGCCAAATCTTTTGAAGAGTTCCCTTTCT 347
Qy 70 AspileGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
Db 348 GCCTTGGAGGCTGGACAGAGCCAGCATTTACTGTAAATAATTAAGTCCCTGAGAAAGT 407
Qy 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyLeuArgSerSerLeuSerThr 109
Db 408 GCTTCACATCTCCATGTGAAATATGTACCATGGGTACCTGACGCTCCTTAAGTACT 467
Qy 110 GlnValIleProAlaIleTyIleLeu 118
Db 468 AAAGTGATACCTGCCATCTACCTCCTG 494

RESULT 36

AJ451155 798 bp mRNA linear EST 19-APR-2002
LOCUS
DEFINITION AJ451155 riken1 Gallus gallus cdna clone 27m8r1, mRNA sequence.
ACCESSION AJ451155
VERSION AJ451155.1 GI:20218376

KEYWORDS

EST.

SOURCE

ORGANISM

Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 798)
Buerstedde, J.M.

Gallus gallus bursal lymphocyte EST

Unpublished (2002)

Contact: Buerstedde JM

Cellular Immunology

Heinrich-Pette-Institut

Martinistr. 52, 20251 Hamburg, Germany

Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.

Location/Qualifiers

1. .798

/organism="Gallus gallus"

FEATURES

source

1. .798

/db_xref="taxon:9031"
/clone="27m8r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/note="CB inbred strain"
BASE COUNT 187 a 170 c 199 g 236 t
ORIGIN

Alignment Scores:

Pred. No.: 5e-25 Length: 798
Score: 360.50 Matches: 83
Percent Similarity: 51.14% Conservative: 52
Best Local Similarity: 31.44% Mismatches: 88
Query Match: 16.88% Indels: 41
DB: 9 Gaps: 6

US-09-208-629F-3 (1-407) x AJ451155 (1-798)

Qy 10 AlaGlySerGlnAspIleLysMetLys-----IleLeuIleLeuVal 23
Db 37 GCCGAGCGGAGGAGGAGGAGGATGCCGGCGCCGGCGCTGCTGCTGCTGCT 96
Qy 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 97 AGCGCCCTGCTGGCG-----GTGGATGACTTGTGGCCAGAAAGTCCCT 111
Qy 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
Db 112 ---GCTCGCTCCACAGAGACAGACAGTGTGGTGGTGGTGGTGGTGGTGGT 147
Qy 64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAla 83
Db 148 -----TCAAAGAGAGAGAGTGTGGTGGCCAGAGGTCCAGATACTAGT 192
Qy 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyLeu 103
Db 193 AACGCTCTGAGGAGTATACAAA-----GTGGATGACTTGTGGCCAGAAAGTCCCT 243
Qy 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyIleLeuLeuPheValValGly 123
Db 244 ACAGAAATCAATACATACAGATTTTCTCCACTGTCTATGTCTTCTTATCATGTGT 303
Qy 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
Db 304 TTGCCAAGCAATGCTATGCCCATCTGGGCTCTTTTTCAGAACAGAGAAACATCCT 363
Qy 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 364 GCTGTGATTATATGTTAACTTGGCATTTGGCAGACCTTCTCTGTTGTTGCCA 423
Qy 163 PheLysIleAlaTyHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 424 CTGAAGATTTGCATATCATTTAAATGCAATAAATGCTGTTTGGGGAAGGTCTCTGCAA 483
Qy 183 IleThrThrValValPheTyTrpGlyAsnMetTyTrpCysAlaIleLeuLeuThrCysMet 202
Db 484 GTGCTGTGGATTTTATGGAATATGTAATGTAATGTAATGTAATGTAATGTAATG 543
Qy 203 GlyIleAsnArgTyLeuAlaThrAlaHisProPheThrTyTrpGlnLysLeuProLysArg 222
Db 544 AGTGTGCAAGGATTTGGGNTAGTGAACCCCATAGTGCACCTCAAGAANG---AAGTCT 600
Qy 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyTrpMetLeuPro 242
Db 601 GAAATGCGCTGGGATCTTNCCTTGTCTGATGACTGATTTGTGGTGAACCATTCG 660
Qy 243 PheValIleLeuLysGlnGluTyHisLeuValHis-SerGluIleThrThr-CysHisA 262
Db 661 CTGTATCTTGTCAATCANACNGNGTATATTTCAGACCTTTAATCACTACCTGCGCATG 720
Qy 262 spValVal 264
|||||


```

QY 10 AlaGlySerGlnAspIleLysMetLys-----IleLeuIleLeuVal 23
Db 39 GCCGGAGCGGAGGAGGAGGATGGCGGGCGCGGCTGTCGTCTGCTGCTT 98
QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 99 AGCGCCCTGCTGGCG-----
QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
Db 114 ---GCTGCCGCTCCACAGACAGTAGAACAGCAGT-----
QY 64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrIleLysAla 83
Db 150 -----TCAAAAGGAAGAGTTTGTGGCCAGAGGTTCCAGATACAGT 194
QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTrpLeu 103
Db 195 AACGCCCTCGAGGAGTCATCAAA-----GTGGATGACTTTCAGCAAAAGTCCTT 245
QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrlleLeuLeuPheValValGly 123
Db 246 ACAGGAAATACTACAGTTTTTCTCCACCTGCTATGTCATTGCTTTATCATTTGTT 305
QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSer 142
Db 306 TTGCCAAGCAATGCTATGGCCATCTGGTCTTTTTCAGAACAAAGAAACATCCT 365
QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
Db 366 GCTGTGATTATATGTTAACTTGGCATTGGCAGACCTCTCTCTGTTGTTGTTCCCA 425
QY 163 PheLysIleAlaTyrlleHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
Db 426 CTGAAGATTGCATCATTTAAATGGCAATACTGCTGTTGGGAAGTCTCTCCAAA 485
QY 183 IleThrThrValValPheTyrlleGlyAsnMetTyrlleCysAlaIleLeuLeuThrCysMet 202
Db 486 GTGCTGTTGGATTTTATGGAATAAGTATGCTGCTCCATCTTTTATGACATGCTC 545
QY 203 GlyIleAsnArgTyrlleAlaThrAlaHisProPheThrTyrlleGlnLysLeuProLysArg 222
Db 546 AATGTGCAAGGATTTGGGTTGTAGTGAACCCCATAGTGCACCTCAAGA---AGGAAGTCT 602
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrlleMetLeuPro 242
Db 603 GAAATTGCCCTGGGCTCTCCCTGCTATCTGATACACTGATTGTTGGGAACCATTCGG 662
QY 243 PheValIleLeuLysGln 248
Db 663 CTGTATCTTGCTCAATCAG 680

RESULT 39
AJ447835
LOCUS      724 bp      mRNA      linear      EST 19-APR-2002
DEFINITION AJ447835 riken1 Gallus gallus cDNA clone 17n1lr1, mRNA sequence.
ACCESSION  AJ447835
VERSION     AJ447835.1 GI:20215056
KEYWORDS    EST.
SOURCE      chicken.
ORGANISM    Gallus gallus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
AUTHORS     Buerstedde, J.M.
TITLE       Gallus gallus bursal lymphocyte EST
JOURNAL     Unpublished (2002)
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institut
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

```

```

FEATURES             Location/Qualifiers
     source            1..724
     db_xref="taxon:9031"
     clone_lib="riken1"
     cell_type="bursal lymphocyte"
     dev_stage="2-3 weeks old"
     note="CB inbred strain"
BASE COUNT  168 a  155 c  190 g  210 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      3,35e-24      Length:      724
Score:          351.50      Matches:      80
Percent Similarity: 50.78%      Conservative: 51
Best Local Similarity: 31.01%      Mismatches: 87
Query Match:    16.46%      Indels:      40
DB:             9      Gaps:      6

US-09-208-629F-3 (1-407) x AJ447835 (1-724)
QY 10 AlaGlySerGlnAspIleLysMetLys-----IleLeuIleLeuVal 23
Db 37 GCCGGAGCGGAGGAGGAGGATGGCGGGCGCGGCTGTCGTCTGCTGCTT 96
QY 24 AlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAsp 43
Db 97 AGCGCCCTGCTGGCG-----
QY 44 AsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPhe 63
Db 112 ---GCTGCCGCTCCACAGACAGTAGAACAGCAGT-----
QY 64 GluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrIleLysAla 83
Db 148 -----TCAAAAGGAAGAGTTTGTGGCCAGAGGTTCCAGATACAGT 192
QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTrpLeu 103
Db 193 AACGCCCTGAGGAGTGCATCAAA-----GTGGATGACTTTCAGCAAAAGTCCTT 243
QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrlleLeuLeuPheValValGly 123
Db 244 ACAGGAAACATACATACAGTTTTTCTCCACCTGCTATGTCATTGCTTTATCATTTG 303
QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSe 142
Db 304 TTGCCAAGCAATGCTATGGCCATCTGGGTCTTTTTTTTCAGAACAAAGAAACATCC 363
QY 142 rLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPr 162
Db 364 TGCTGTGATTATATGTTAACTTGGCATTGGCAGACCTTCTCTGTTGCTGTTGCC 423
QY 162 oPheLysIleAlaTyrlleHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysAr 182
Db 424 ACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTGGGAAGGCTCTGCAA 483
QY 182 gIleThrThrValValPheTyrlleGlyAsnMetTyrlleCysAlaIleLeuLeuThrCysMe 202
Db 484 ACTGCTGTTGTAATTTTTTATGGAATATATGACTGCTCCATTTCTTTTATGACATGCT 543
QY 202 tGlyIleAsnArgTyrlleLeuAlaThrAlaHisProPheThrTyrlleGlnLysLeuProLysAr 222
Db 544 CAGTGTGCAAGGATATGGTGTGATGAACCCCATAGTGCACCTCAAGA---AGGAAGTC 600
QY 222 gSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrlleMetLeuPr 242
Db 601 TGAAATTGCCCTGGGCATCTCCCTGCTATCTGGATACATGATTGTTGTTGGGAACCATTC 660
QY 242 oPheValIleLeuLysGlnTyrlleHisLeuValHisSerGluIleThrThr 259
Db 661 GCTGTATCTTGTCAATTCAGACNGGGTATATTTTCAGACCTTAACATCACTAC 712

```

RESULT 40
 CNS03DLV/c
 LOCUS
 DEFINITION
 Tetraodon nigroviridis genome survey sequence T7 end of clone
 017G16 of library G from Tetraodon nigroviridis, genomic survey
 sequence.

ACCESSION
 AL239260
 VERSION
 AL239260.1 GI:7898395
 KEYWORDS
 GSS; genome survey sequence.
 SOURCE
 Tetraodon nigroviridis.
 ORGANISM
 Tetraodon nigroviridis.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neoceratrygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE
 1 (bases 1 to 817)
 Roest-Crollius, H., Jaillon, O., Dasilva, C., Bounneau, L., Fisher, C.,
 Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
 Saurin, W., and Weissenbach, J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence

TITLE
 Tetraodon nigroviridis DNA sequence

JOURNAL
 Unpublished

REFERENCE
 2 (bases 1 to 817)
 Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.,
 Bounneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
 Weissenbach, J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis.

TITLE
 Tetraodon nigroviridis DNA sequence

JOURNAL
 Unpublished

REFERENCE
 3 (bases 1 to 817)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000)

TITLE
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.

FEATURES

Source
 1..817
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="017G16"
 /clone_lib="G"
 /note="Genoscope sequence ID : COBG017BD08LP1-end : T7"
 BASE COUNT 218 a 183 c 232 g 167 t 17 others
 ORIGIN

Alignment Scores:

Pred. No.: 3.9e-24 Length: 817
 Score: 351.50 Matches: 71
 Percent Similarity: 59.52% Conservative: 29
 Best Local Similarity: 42.26% Mismatches: 58
 Query Match: 16.46% Indels: 10
 DB: 17 Gaps: 3

US-09-208-629f-3 (1-407) x CNS03DLV (1-817)

QY 101 GlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIle-TyrIleLeuLeuPh 120
 Db 521 GGAACCTCTTCGCGCGCGCTCCACCGCTTCTTCCCGCTGCTATACATCACCGTCTT 462
 QY 120 eValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr--Ly 139
 Db 461 CGTGGTGGGCTCCCAACCAACCCCTGGCCATCTGGGCTTCTCTTCGCGCAAGAA 402
 QY 139 sSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysVa 159
 Db 401 GAAGCATCCATCTCCATCTTATGGCCAACTGGCGCTGGCGGACCTGCTCTTCGTCAT 342
 QY 159 lThrLeuProPheIysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVa 179
 Db 341 CTGGGTCCCGCTGAAATCGCCTACCACTCAACGGGACGACTGGGTCTACGCGGAGTC 282

QY 179 lMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLe 199
 Db 281 CCTGTGCAAGGTCTCTGTGGCTTCTTCTACGCAACATGTACTGTCTCCATCGCCTTCAT 222
 QY 199 uThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysIle 219
 Db 221 CCGGTGCATCAGCGCTCCAGCGCTGCAAGCGCTGTCTCCACCG-CT 177
 QY 219 uProLysArgSerPheSerLeuLeuMetCysGlyIle-----ValTrpValMetVa 236
 Db 176 CGACAGAGGGGGGCGAGCGGTGCGCGCTGGCGGTGTCGCGGCCATCTGCGTGTGGT 117
 QY 236 lPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGl 256
 Db 116 GTGGCTCATCACCGTTCCTCTGTACTGTACGATCAGCAGGTCTCTGGTGACAAACCTGGG 57
 QY 256 uIleThrThrCysHisAspVal 263
 Db 56 CATCCGCACTGCCACGACGTC 35

Search completed: June 29, 2003, 10:14:30
 Job time : 1227.52 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 09:48:59 ; Search time 131.838 Seconds
(without alignments)
4582.672 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLTXXHPVAGSODIKMKIL.....AMARPLXRPRIWDIHAH 407

Scoring table:
BLOSUM62 Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1055720 seqs, 742224136 residues
Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09208629/runat_24062003_101628_13027/app.query.fasta_1.1166
-DB=Published Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09208629 @CEN 1.153 @runat_24062003_101628_13027
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCX=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1443.5	67.6	1830	9 US-10-225-567A-301	Sequence 301, App
2	563	26.4	1451	9 US-10-225-567A-299	Sequence 299, App
3	509	23.8	3592	9 US-10-225-567A-323	Sequence 323, App
4	508	23.8	3299	9 US-10-208-408-24	Sequence 24, Appl

5	507.5	23.8	4895	9 US-10-225-567A-515	Sequence 515, App
6	500	23.4	1425	9 US-10-094-417-9	Sequence 9, Appl
7	500	23.4	2137	9 US-09-782-974C-75	Sequence 75, Appl
8	500	23.4	2137	9 US-10-225-567A-529	Sequence 529, App
9	496	23.2	1955	9 US-10-190-469-2	Sequence 2, Appl
10	493	23.1	1080	10 US-09-739-151-1	Sequence 1, Appl
11	471.5	22.1	2588	10 US-09-943-718-3	Sequence 3, Appl
12	458	21.4	1080	10 US-09-943-718-5	Sequence 5, Appl
13	431.5	20.2	1086	9 US-10-222-024-1	Sequence 1, Appl
14	431.5	20.2	1086	9 US-10-251-385-77	Sequence 77, Appl
15	431.5	20.2	1638	9 US-10-225-567A-111	Sequence 111, App
16	428.5	20.1	1086	9 US-10-251-385-205	Sequence 205, App
17	398.5	18.7	2070	9 US-10-225-567A-303	Sequence 303, App
18	395	18.5	1020	9 US-10-251-385-31	Sequence 31, Appl
19	389	18.2	1020	10 US-09-788-133-1	Sequence 1, Appl
20	389	18.2	2299	9 US-10-251-385-181	Sequence 181, App
21	388.5	17.7	2299	9 US-10-225-567A-224	Sequence 224, App
22	378.5	17.7	993	10 US-09-826-791-1	Sequence 1, Appl
23	378.5	17.7	1041	9 US-09-828-478-1	Sequence 1, Appl
24	378.5	17.7	1041	10 US-09-826-791-5	Sequence 5, Appl
25	378.5	17.7	1041	10 US-09-866-230-6	Sequence 6, Appl
26	378.5	17.7	1260	9 US-09-779-679-1	Sequence 1, Appl
27	378.5	17.7	1260	9 US-09-779-679-24	Sequence 24, Appl
28	378.5	17.7	1430	9 US-09-828-478-3	Sequence 3, Appl
29	378.5	17.7	1700	10 US-09-728-952-26	Sequence 26, Appl
30	378.5	17.7	2807	9 US-10-225-567A-588	Sequence 588, App
31	378	17.7	1014	9 US-10-225-567A-546	Sequence 546, App
32	376.5	17.6	2424	9 US-10-225-567A-218	Sequence 218, App
33	374	17.5	1358	9 US-10-167-192-4	Sequence 4, Appl
34	369.5	17.3	1014	9 US-10-023-775B-1	Sequence 1, Appl
35	369.5	17.3	1014	9 US-10-270-144-1	Sequence 1, Appl
36	369.5	17.3	1014	9 US-10-188-405-7	Sequence 7, Appl
37	369.5	17.3	1014	9 US-09-885-453-2	Sequence 2, Appl
38	369.5	17.3	1014	9 US-10-079-384-13	Sequence 13, Appl
39	369.5	17.3	1014	9 US-10-225-567A-646	Sequence 646, App
40	369.5	17.3	1014	10 US-09-943-798-3	Sequence 3, Appl
41	368	17.2	1301	9 US-10-024-494-7	Sequence 7, Appl
42	367	17.2	1146	9 US-10-225-567A-413	Sequence 413, App
43	366	17.1	1429	9 US-09-077-173A-1	Sequence 1, Appl
44	362.5	17.0	1098	9 US-10-225-567A-331	Sequence 331, App

ALIGNMENTS

RESULT 1
US-10-225-567A-301
; Sequence 301, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 301
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-301

Alignment Scores:
Pred. No.: 1,82e-125 Length: 1830
Score: 1443.5 Matches: 264
Percent Similarity: 83.74% Conservative: 45
Best Local Similarity: 71.54% Mismatches: 59


```
QY 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
DB 589 TGAATGCTGCTATTGCTTTCTATGCAACATGTAAGTCTTCTTCAATGACC 648
QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
DB 649 TGCCTAGTGTGAGAGATATGGGTGATCGTGAACCCATGGGCACTCCAGG---AAG 705
QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
DB 706 AAGCAAAACATGCGATGCGATCTCCCTGCAATATGGTGTGATTCGTGCTGGTCACC 765
QY 241 LeuProPheValIleLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
DB 766 ATCCCTTTGATGTGCTGGAAGCAGCACCATTCTTCTCCCTGGAACATCAGCACCTGT 825
QY 261 HisAspValValAspAlaCysGluSerProSerSerPhe-----ArgPheTyr 276
DB 826 CATGATGTTTIG-----CCTGAGCAGCTCTTGTGGGAGACATGTTCAAT 870
QY 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCys 296
DB 871 TACTTCCTCTCTGGCCATGCGGTCTTCTGTTCACGCTTCTCCAGCCTCTGCC 930
QY 297 TyrThrThrLeuLeuHisLysLeuLysSer-----LysAspArgIle 310
DB 931 TATGTGCTGATGATCAGAACTGCTGCGATCTTCTGCCATGATGAACACTCAGAGAAGAA 990
QY 311 TrpLeuGlyTyrIleLysAlaValLeuLeuLeuIleValPheThrIleCysPheAla 330
DB 991 AGGAAGAGGCCATCAAACTCATGTGCTGCTGCGCCATGATCACTGATCTGCTCACT 1050
QY 331 ProThrAsnIleLeuLeuIleHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
DB 1051 CCTAGTAACCTCTGCTGTGGTGCAATATTCTGATTAGAGCCAGGCCAGCCAT 1110
QY 351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
DB 1111 GTCTATGCCCTGTACATGTAGCCCTCTGCTCTCTACCCCTTAACAGCTGCATCGACCC 1170
QY 371 PheLeuTyrPheValMetSer 377
DB 1171 TTGTCTATTACTTTGTTTCA 1191

RESULT 3
US-10-225-567A-323
; Sequence 323, Application US/10225567A
; Publication No. US2003013798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 323
; LENGTH: 3592
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-323

Alignment Scores:
Pred. No.: 1,32e-37 Length: 3592
Score: 509.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.83% Indels: 42
```

```
DB: 9 Gaps: 14
US-09-208-629F-3 (1-407) x US-10-225-567A-323 (1-3592)
QY 20 LeuIleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
DB 360 CTGCTCTGCTGGTGGCCGCTGCTTC-----AGTCTGTGGCGCCGCTGTTG 404
QY 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
DB 405 TCTGCCCGCACCAGCCGCGCCAGCCAGCAATCAAAAGCAACAAATCCACCTTAGATCCC 464
QY 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
DB 465 CGGTCAATTTCTTCTCAGGAACCCCAATGATGAATGAACCAATTTTGGGAGGATGAGGAG 524
QY 72 GluGlyTyrThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87
DB 525 AAAAATGAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAGCAGCTCCTCTT 584
QY 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
DB 585 CAAAAACAACCTTCTCTGCATTCATCTCAGAGATGCCCTCGGATATTTGACCAGCTCTCTGG 644
QY 108 SerThrGlnValIleProAlaIleIleLeuLeuPheValValGlyValProSerAsn 127
DB 645 CTGACACTCTTTGTCCTCATCTGTACACCGGAGTGTGTTGTCTGCTGCTCCCTTTAAGATC 704
QY 128 Ile-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
DB 705 ATCATGCCATCGTGTGTTTCATCTGAAATGAAGTCAAGAAG---CGGCGGTGGTG 761
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
DB 762 TACATGCTGCACCTGCGCAGCAGATGTGCTGTTGTCTGCTGCTCCCTTTAAGATC 821
QY 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThr 185
DB 822 AGCTATTACTTTTCCGCGCAGTATTGGCAGTTTGGCTCTGAATTGTGCTGCTCACT 881
QY 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsn 205
DB 882 GCAGCATTTTACTGTAAACATGTACGCCCTCTATCTTCTCATGACAGTCATCAAGCATGAC 941
QY 206 ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg 222
DB 942 CGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCTGGCGTACTCTGGAAAGGCT 1001
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
DB 1002 TCCTTCACT-----TGTCTGGCCATCTGGGCTTTGGCCATCGCAGGGTAGTGCCT 1052
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
DB 1053 CTGCTCCTCAAGGAGCAACCACTCCAGGTGCCGCGCTCAACATCACTACCTCTCATGAT 1112
QY 263 ValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAla 282
DB 1113 GTGCTCAATGAACCCCTGCTCGAAGGCTACTATGCTCTACTTCTCAGCCTTCTCTGCT 1172
QY 283 PhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHis 302
DB 1173 GTCTTC--TTTGTGCGCGTGTATCTCCACGGTCTGTTATGTGCTATCATTCGA 1229
QY 303 LysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
DB 1230 TGTCTTAGCTCTTCCGAGTTCGCAACCGCAGCAAGAGTCCGCGCTTTG-----TTC 1283
QY 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
DB 1284 CTGTCAGCTGCT-----GTTTCTGCATCTTCTCATTTGTTTCCGACCCCAACAGTC 1337
QY 335 IleLeuValIleHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr 352
```

```
Db 1338 CTCCTGATGGCATTAC--TCATTCTCTTCTCACATTCACAGAGGCTGCCTAC 1394
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 353 PheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu 372
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1395 TTTGCCCTACCTCTCTGTGTGTGTGCAGCAGCATAAGCTCGTGCATCGACCCCTAAAT 1454
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 373 TyrPheValMetSer 377
      ::::: :::::
Db 1455 TACTATTACGCTTCC 1469

RESULT 4
US-10-208-408-24
; Sequence 24, Application US/10208408
; Publication No. US20030096272A1
; GENERAL INFORMATION:
; APPLICANT: Schebye, Xiao Min
; TITLE OF INVENTION: GENES REGULATED BY PEROXISOME PROLIFERATOR-ACTIVATED RECEPTOR
; FILE REFERENCE: PA-0048-1 US
; CURRENT APPLICATION NUMBER: US/10/208,408
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,868
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 3299
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030096272A1 2456481CB1
US-10-208-408-24

Alignment Scores:
Pred. No.: 1.45e-37 Length: 3299
Score: 508.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.78% Indels: 42
DB: 9 Gaps: 14

US-09-208-629F-3 (1-407) x US-10-208-408-24 (1-3299)
Qy 20 LeuIleValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 74 CTGCTGTGGTGGCGCGCTGCTTC-----AGTCTGTGGCGCGCGCTGTG 118
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
      ::::: :::::
Db 119 TCTGCCCGCACC CGCGCGCGCAGCCAGCAATCAAAAGCAAAATGCCACCTTAGATCCC 178
      ::::: :::::
Qy 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
      ::::: :::::
Db 179 CGGTCACTTTCTCTTCAGGAACCCCAATATAATATGAACCAATTTGGGAGGATGAGGAG 238
      ::::: :::::
Qy 72 GluclyIrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87
      ::::: :::::
Db 239 AAAAATGAAGTGGGTGTTAACTGAATACAGATTAGTCCATCAATAAAACAGCTCTCTT 298
      ::::: :::::
Qy 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
      ::::: :::::
Db 299 CAAAAACAACCTCTCTGATTCATCTCAGAGATCCCTCCGGATATTTGACCAGCTCTCTGG 358
      ::::: :::::
Qy 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
      ::::: :::::
Db 359 CTGACACTTTTGTGCCATCTGTGTACACCGGAGTGTGTGTAGTCAGCTCCCACTAAAC 418
      ::::: :::::
Qy 128 Ile-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
      ::::: :::::
Db 419 ATCATGCCCATCGTGTGTGTTCTCATCTCTGAAATGAAGTCAAGAG-----CCGGCGGTGGT 475
      ::::: :::::
Qy 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
```

```
Db 476 TACATGCTGCACCTGCCACGCGCAGATGCTGTTTGTGTGTGCTCCCTTTAAGATC 535
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 185
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 536 AGCTATTACTTTTCCGGCAGTGAATGGCGATTTGGGTCTGAATTTGTGCTGCTCACT 595
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsn 205
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 596 GCAGCATTTTACTGTAACATGACGCTCTATCTTGCTCTGACAGTCAAGCATGAC 655
      ::::: :::::
Qy 206 ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg 222
      ::::: :::::
Db 656 CGGTTTCTGGCTGTGGTGATCCCATGCACTCCCTCTCTGGGTACTCTGGGAAGGGCT 715
      ::::: :::::
Qy 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
      ::::: :::::
Db 716 TCTTTCAC-----TGTCTGGCCATCTGGGCTTGGCCATCGCAGGGTAGTGCCT 766
      ::::: :::::
Qy 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
      ::::: :::::
Db 767 CTCGTCCTCAAGAGCAACCATCCAGGTCGGGCTCAACATCACTACTCTGCTCATGAT 826
      ::::: :::::
Qy 263 ValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeuAla 282
      ::::: :::::
Db 827 GTGCTCAATGAAACCCCTGCTCGAAGGCTACTATCCCTACTACTTCTCAGCCTTCTCTGCT 886
      ::::: :::::
Qy 283 PhePheGlyPheLeuLeuProPheValIleIlePheCysTyrThrThrLeuIleHis 302
      ::::: :::::
Db 887 GTCTTC---TTTTTGTGCGCTGATCAATTCACCGTCTGTATGTCTATCATTCGA 943
      ::::: :::::
Qy 303 LysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
      ::::: :::::
Db 944 TGTCTTAGCTCTCCGCACTTCCCAACCCGACAGCAAGTCCCGGCTTGT-----TTC 997
      ::::: :::::
Qy 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
      ::::: :::::
Db 998 CTGTCACTGCT-----GTTTTCTGCATCTTTCATCATTTGCTTCGGACCCACAAAGTC 1051
      ::::: :::::
Qy 335 IleLeuValIleHisHisAlaAsnTyrTyrTyrHis-----AsnThrAspSerLeuTyr 352
      ::::: :::::
Db 1052 CTCCTGATTGCGCATTCAC---TCATTCTTCTTCACACTTCCACCCACAGGCTGCCTAC 1108
      ::::: :::::
Qy 353 PheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu 372
      ::::: :::::
Db 1109 TTTGCCCTACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1168
      ::::: :::::
Qy 373 TyrPheValMetSer 377
      :::::
Db 1169 TACTATTACGCTTCC 1183

RESULT 5
US-10-225-567A-515
; Sequence 515, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 515
; LENGTH: 4895
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-515
```

Alignment Scores:

Pred. No.:	2,86e-37	Length:	4895
Score:	507.50	Matches:	111
Percent Similarity:	52.97%	Conservative:	76
Best Local Similarity:	31.44%	Mismatches:	151
Query Match:	23.76%	Indels:	15
DB:	9	Gaps:	6

US-09-208-629F-3 (1-407) x US-10-225-567A-515 (1-4895)

```

QY   27 LeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSera 46
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   188 CTGCTCTGTGGCCCTGGTGTG-----GGTTACAGCTGTGGCGGCACCAG 238

QY   47 LysProThrLeuThrIleLys-----SerPheAenGlyGlyProGlnAsnThrPheGlu 64
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   239 ACCCCAGCGCTACGACGAGCGGAGCACCGGAGGTGTATGACAGCAGCCCTCA 298

QY   65 GluPheProLeuSerAspileGluGlyTrpThrGlyAlaThrThrIleLysAlaGlu 84
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   299 ATCTGCCCTGCCCCC-----CGCGGCTACCCAGGCCAAAGTC----- 334

QY   85 CysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrrLeuArg 104
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   335 TGTGCCAATGACAGT---GACACCTTGAGCTCCCGACAGCTCAGCGGACTGCTTCG 391

QY   105 SerSerLeuSerThrClnValleProalaleTyrrilleLeuLeuPheValValGlyVal 124
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   392 GGTGGTGTCCCCACCAGGCTGTGCCCGCCCTCTATGGCTGTGCTGTGGTGGGCTG 451

QY   125 ProSerAsnilevalThrTrpLysLeuSerLeuArgThrLysSerIleSerLeuVal 144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   452 CGGCCAATGGGCTGGCGTGTGGTGTGGCCAGCAGGACCTCGGCTGCCCTCCACC 511

QY   145 IlePheHieThrAsnLeuAlaIleAlaaspLeuLeuPheCysValThrLeuProPhelys 164
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   512 ATGCTGCTGATGAACCTCGGACTGTGACCTCTGCTGGCCCTGGCGCTGCCCGCGG 571

QY   165 IleAlatyHisLeuAnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   572 ATCGCTACACCTCGTGGCCAGGCTGGCCCTTCGGGGAGCGCGCTGGCGCTGCC 631

QY   185 ThrValValPheTyrrGlyAsnMetTyrrCysAlaIleLeuLeuThrCysMetGlyIle 204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   632 ACGCGCGACTCATGGTCACATGATGCTCAGTGTGCTGTGGCGCGCGCTCAGGCTG 691

QY   205 AsnArgTyrrLeuAlaThrAlahisProPheThrTyrrGlnLysLeuProLyAsrSerPhe 224
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   692 GATCGTACTGCGCCCTGTGCACCGCTCGCGGCGCGCGCTCGGCGCGCGCTG 751

QY   225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrrMetLeuProPheVal 244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   752 GCCCTGGACTTCGATGCTGCTGTGGCTCATGGCGCGCGCGCTGCCCTCCAGCA 811

QY   245 IleLeuLysGlnTyrrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   812 CTGACGGCGAGACCTTCGGCTGGCGCGCTCGATCGCTGCTGCGATGACGCGCTG 871

QY   265 AspAlaCysGluSerProSerSerPheArgPheTyrrTyrrPheValSerLeuAlaPhePhe 284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   872 ---CCCTGGAGCAGCAGCGCTCCCACTGGCAACCGGCTTCACCTGCTGGCGCTGTG 928

QY   285 GlyPheLeuIleProPheValIlellePheCystTyrrThrThrLeuIleHisLysLeu 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   929 GGCTGTTCTGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 988

QY   305 LysSerLysAspArgIleTyrrLeuGlyTyrrilleLysAlaValLeuLeuValIle 324
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   989 GCGCGCAGCGCGCGCTACGGCCACCGCTGAGGCTGACCGCAGTGGTCTGGCTCC 1048

QY   325 PheThrIleCysPheAlaProThrAsnIlellelleValIleHisAlaAsnTyrr 344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```
Db 469 CTTTCAAAATCTACTACCTTGAACACCGCCACCTGGGTATTCGGGGTGTCTGCTTTC 528
Qy 182 ArgileThrThrValValPheTyrGlyAsnMetTyrCysAlaLeuLeuLeuLeuThrCys 201
Db 529 AACGTGTGACCGGCGCTTTACGAAACATGATTCACAGATCCACACATGACCTGT 588
Qy 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys 221
Db 589 ATCAGCGTGAGCGCTTCTGGGGTCTGTACCGCTCAGCTCCAGCGCTGGCGCGC 648
Qy 222 ArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeu 241
Db 649 CGTCTTACCGGCGCGGCTGTGCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 708
Qy 242 ProPheValLeuLeuLysGlnGluTyrHisLeuValHisSerGluLeuThrThrCysHis 261
Db 709 CCGCTGGCGCGCACCGATCTACCTACCGGTGCGACCGCTGGGCATCATCCTGCTTC 768
Qy 262 AspValValAspAlaCysGluSerProSer-----SerPheArgPheTyr 276
Db 769 GACGTCTCAAGTGCAGATGCTCCCGAGCGTGGCCATGTGGCGCTGTTCTCTTCACC 828
Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuLeuProPheValLeuLeuLeuPheCys 296
Db 829 ATCTTCATCTGCTG-----TCTCATCCCGTTCGTGATCATCCGTGGCTGT 876
Qy 297 TyrThrThrLeuLeuHisLysLeu-----LysSerLysAsp 308
Db 877 TACAGCGCCACCATCTCAAGCTGTGGCAGCGGAGCGACGCGCGGAGCAGCG 936
Qy 309 ArgileThrLeuGlyTyrIleLysAlaValLeuLeuLeuValLeuThrThrCys 328
Db 937 AGCGCGCGGTGGCGCTGGCGCGGTGCTTGTGCTG-----GCCTTGTACCTGTC 987
Qy 329 PheAlaProThrAsnIleLeuValLeuValLeuHisAlaAsnTyrTyrHisAsnThr 348
Db 988 TTCGCCCCCAACAACCTTCTGCTCTGCGCGCACATCGTGAGCGCGCTGTCTACGGC-- 1044
Qy 349 AspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 1045 AAGAGCTACTACCACTGTCAACGTCACGCTGTGCTCAGCTGCTCCTCAACAACCTGT 1104
Qy 369 AspProPheLeuTyrPheValMetSerLys 378
Db 1105 GACCGGTTGTTTATTACTTGTGCTCCGG 1134
```

RESULT 7

```
US-09-782-974C-75
; Sequence 75, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534A1e1 G Protein Coupled Receptor
; FILE REFERENCE: 411USPHR311
; CURRENT APPLICATION NUMBER: US/09/782.974C
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165,838
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/198,568
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,071
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,678
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/173,396
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,129
```

```
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/185,421
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,554
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,530
; PRIOR FILING DATE: 2000-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: Patent version 3.1
; SEQ ID NO 75
; LENGTH: 2137
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-974C-75

Alignment Scores:
Pred. No.: 4,34e-37 Length: 2137
Score: 500.00 Matches: 113
Percent Similarity: 52.12% Conservative: 59
Best Local Similarity: 34.24% Mismatches: 128
Query Match: 23.41% Indels: 30
DB: 9 Gaps: 8

US-09-208-629F-3 (1-407) x US-09-782-974C-75 (1-2137)

Qy 69 SerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAsp 88
Db 150 TCTAGCTCTACCTGCTACTTCTGCGCTGCTTCT-----GCACAGACCCGGCGGAG 203
Qy 89 SerIleSerThrLeuHisVal-----AsnAsnAlaThrIleGlyTyr 102
Db 204 GACCCCTCCAGGATGAGGTCCGAAACAGCACCGCGCCGACACCGCGTGCAGATG 263
Qy 103 LeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 122
Db 264 CTCGGAACCGCGCATCGCGTGGCCCTGCGCGCTGCTACTCGCTGTGGCGCGGTC 323
Qy 123 GlyValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSer 142
Db 324 AGCATCCCGGCACTCTTCTCTGTGGTGTCTGTCGGCGCATGGGGCCACATCC 383
Qy 143 ---LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeu 161
Db 384 CCGTGGTCTATCTTCATGATCAACCTGAGCGTCACGACCTGATGCTGGCCAGCGGTG 443
Qy 162 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCys 181
Db 444 CTTTCCAAATCTACTACCATTTGCAACCGCCACCTGGGTATTCGGGGTGTCTTTC 503
Qy 182 ArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCys 201
Db 504 AAGTGTGACCGTGGCGCTTTTACGCAACATGATTCAGCAATCTTCACCATGACCTGT 563
Qy 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys 221
Db 564 ATCAGCGTGAGCGCTTCTGGGGTCTGTACCGCTCAGCTCCAGCGCTGGCGCGC 623
Qy 222 ArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeu 241
Db 624 CGTCTTACCGGCGCGGCTGTGCAGGAGACCTGCTGCTGCTGCTGCTGCTGCTGCT 683
Qy 242 ProPheValLeuLysGlnGluTyrHisLeuValHisSerGluLeuThrThrCysHis 261
Db 684 CCGCTGGCGCGCACCGATCTACCTACCGGTGCGACCGCTGGGCATCATCCTGCTTC 743
Qy 262 AspValValAspAlaCysGluSerProSer-----SerPheArgPheTyr 276
Db 744 GACGTCTCAAGTGCAGATGCTCCCGAGCGTGGCCATGTGGCGCTGTTCTCTTCACC 803
Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuLeuProPheValLeuLeuLeuPheCys 296
Db 804 ATCTTCATCTGCTG-----TTCCTCATCCCGTTCGTGATCATCCGTGGCTGT 851
```

162	Qy	ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCys	181
444	Db	CCTTTCCAAATCTACTACCATTCGACCGCCACCTGGGTATTTCGGGGTGTGCTTGC	503
182	Qy	ArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCys	201
504	Db	AACGTGGTAGCGTGGCCCTTTACGMAACATGATTCCAGCATCTCTCCACCTAGCACTGT	563
202	Qy	MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys	221
564	Db	ATCAGCGTGAGCGCTTCCTGGGGTCTCTATCCCGCTCAGCTCAAAGCGCTGGCGCGCG	623
222	Qy	ArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeu	241
624	Db	CGTCGTACCGGTGGCGCGGTGTGCAGGACCTGGCTGTCTGTCTTCGACCGCCCTGCC	683
242	Qy	ProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHis	261
684	Db	CCGCTGGCGCGCACCATCTTCACTACCGCGTACGCGCTGGCATCACCATCTGTTCT	743

RESULT 8

744	Db	...	GAGTCTCAAGTGGACGATCTCCACGCGTGGCCATGTGGCGGTTCCTCTTACC	803
277	Qy		TyrPheValSerLeuAlaPheGlyPheLeuIleProPheValIleIleIlePheCys	296
804	Db		ATCTTCATCTCGTG-----TTCTCATCCGTTCTGTGATCATCCGTGGCTGTG	851
297	Qy		TyrThrThrLeuIleHisLysLeu-----LysSerLysAsp	308
852	Db		TACAGGCCACCATCTCAAGCTGTGGCGACGGAGGCGCAGCGCCGGGAGCAGCGG	911
309	Qy		ArgIleTrpLeuGlyTyrlleLysAlaValLeuLeuIleValIlePheThrIleCys	328
912	Db		AGCGCGCGTGGCGCTGGCGCGGTGTCTCTG-----GCCTTTGTCACTGC	962
329	Qy		PheAlaProThrAsnIlelleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThr	348
963	Db		TTCGCCCCAACACTTCGTGCTCGTGGCGCATCGTGTAGCGCCCTGTTCTACGGC---	1019
349	Qy		AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu	368

Alignment Scores:

Qy 369 AspProPheLeuTyrPheValMetSerLys 378
 |||||:::||:::||||:
Dd 1080 GACCCGTTCTATTACTTTGGTCGCCG 1109
 |||||:::||:::||||:
RESULT 9
US-10-190-469-2

; sequence 2, Application US/10190469
: Publication No. US20030013155A1

/ GENERAL INFORMATION:
 / APPLICANT: Glucksmann, Maria A.
 / TITLE OF INVENTION: 14900 Receptor, A No. US20030013155A1el G-Protein Couple
 / FILE REFERENCE: S800-7, 035800/16S196
 / CURRENT APPLICATION NUMBER: US/10/190,469
 / CURRENT FILING DATE: 2002-07-05
 / PRIOR APPLICATION NUMBER: US/09/137,063A
 / PRIOR FILING DATE: 1999-08-20
 / NUMBER OF SEQ ID NOS: 3
 / SOFTWARE: PatentIn Ver. 2.0
 / SEQ ID NO 2
 / LENGTH: 1955
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (1)..(1955)
 / OTHER INFORMATION: n = a, t, c or g
 / US-10-190-469-2

Alignment Scores:

```
Pred. No.: 9,04e-37 Length: 1955
Score: 496.00 Matches: 109
Percent Similarity: 51.99% Conservativeness: 61
Best Local Similarity: 33.33% Mismatches: 133
Query Match: 23.22% Indels: 24
DB: 9 Gaps: 7

US-09-208-629F-3 (1-407) x US-10-190-469-2 (1-1955)

Qy 69 SerAspIleGluGlyThrPheGlyAlaThrThrThrIleLysAlaGluCysProGluAsp 88
Db 222 TCTGAGCTCACTGCTACTTCTGCGCTGCTTCT-----GCACAGACCGCGGCGAG 275
Qy 89 SerIleSerThrLeuHisVal-----AsnAsnAlaThrIleGlyTyr 102
Db 276 GACCCCTCCAGGATGAGGTCGCGAAGACAGACCGCGCGGACACGCGCTGCGAGATG 335
Qy 103 LeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheValVal 122
Db 336 CTGCGGAACCGCGCATGCGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 395
Qy 123 GlyValProSerAsnIleValThrLeuTyrPheLysSerLeuArgThrLysSerIleSer 142
Db 396 AGCATCCCGGCAACCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
Qy 143 ---LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeu 161
Db 456 CGGTCGCTCATCTTCATGATCAACCTGAGCGTCACGAGCATGATGCTGCGCAGCGTGTG 515
Qy 162 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCys 181
Db 516 CTTTCCAAATCTACTACCTATGCAACCGCGCACCATGCTGCTGCTGCTGCTGCTGCTG 575
Qy 182 ArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCys 201
Db 576 AACGTGTGACCGTGGCTTTTACGAAACATGATTTCCAGCATCTCTCACCATGACCTGT 635
Qy 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys 221
Db 636 ATCAGCTGAGCGCTTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 695
Qy 222 ArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeu 241
Db 696 CGTCTGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
Qy 242 ProPheValIleLeuLysGlnIleTyrHisLeuValHisSerGluIleThrThrCysHis 261
Db 756 CCGCTGGCGCGCACCGATCTCACCTACCGGTGCGACCGCTGCGCATCATCATCTGCTTC 815
Qy 262 AspValValAspAlaCysGluSerProSer-----SerPheArgPheTyr 276
Db 816 GACGTCTCAAGTGGAGCATGCTCCCGCGCTGCGCATGCGGCGGCTGCTTCTCTTCCACC 875
Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCys 296
Db 876 ATCTTCATCTGCTG-----TTCTCATCTCGTTCGTGATCATCCGCTGCTGCTGT 923
Qy 297 TyrThrThrLeuIleHisLysLeuLysSerLysAspArgIleTyr----- 311
Db 924 TACACGGCCACCATCTCAAGCTTGTGCGACGGAGGCGGCGCGCGCGGAGCAGCGG 983
Qy 312 LeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIleCysPheAlaPro 331
Db 984 AGCCCGCGGTGGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043
Qy 332 ThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeu 351
Db 1044 AACAACTTCGCTCTCCGGCCACATCGTGAGCGCGCTGCTTCTACGGC-----AAGAGGTAC 1100
Qy 352 TyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 371
Db 1101 TACCACGTGTACAAGCTCACGCTGTGCTCAGCTGCTCAACAACTGCTGCGACCGCTTT 1160
```

```
Qy 372 LeuTyrPheValMetSerLys 378
Db 1161 GTTATTACTTTGCGTCCCGG 1181
```

RESULT 10

```
US-09-739-151-1
; Sequence 1, Application US/09739151
; Patent No. US20010029032A1
; GENERAL INFORMATION:
; APPLICANT: Yuan Zhu
; APPLICANT: Xiaotong Li
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: PAUL, A G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: GP-70567-CL
; CURRENT APPLICATION NUMBER: US/09/739,151
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/413,534
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/103,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-739-151-1
```

```
Alignment Scores:
Pred. No.: 7,32e-37 Length: 1080
Score: 493.00 Matches: 106
Percent Similarity: 53.54% Conservativeness: 53
Best Local Similarity: 35.69% Mismatches: 116
Query Match: 23.08% Indels: 22
DB: 10 Gaps: 6
```

US-09-208-629F-3 (1-407) x US-09-739-151-1 (1-1080)

```
Qy 96 AsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIle 115
Db 28 GACACACGCGATGTCAGATGTCGCGAACCCCGCGATCGCGTGGCCCTGCGCGTGGTG 87
Qy 116 TyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeuSer 135
Db 88 TACTCGCTGGTGGCGCGGTGACATCCCGGCAACCTCTCTCTCTGTTGGGTGCTGTC 147
Qy 136 LeuArgThrLysSerIleSer---LeuValIlePheHisThrAsnLeuAlaIleAlaAsp 154
Db 148 CGGCGCATGGGCGCCAGATCCCGTGGTCACTTTCATGATCACTGAGCGTCAGGAC 207
Qy 155 LeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyr 174
Db 208 CTGATGCTGCCAGCGCTGTTGCTTTCACAAATCTACTACCATTTGCAACCGCCACCTGG 267
Qy 175 ValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCys 194
Db 268 GTATTCGGGTGCTGCTTTCGCAACGTTGACCGTGGCCCTTTTACGCAACATGATATCC 327
Qy 195 AlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPhe 214
Db 328 AGCATCTCACCATGACCTGATCAGCGTGGAGCGCTTCTTGGGGGTCTGTGACCGCTC 387
Qy 215 ThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyrVal 234
Db 388 AGCTCCAAGCGCTGGCGCGCGCTGCTTACCGCGTGGCGCGGTGTCAGGAGCACTGGCTG 447
Qy 235 MetValPheLeuTyrMetLeuProPheValIleLeuLysGlnIleThrHisLeuValHis 254
Db 448 CTGCTCTGACCGCCCTGCTCCCGCGTGGCGCGCACCGATCTCACTACCGGTGCGACGCC 507
Qy 255 SerGluIleThrThrCysHisAspValAlaCysGluSerProSer----- 271
```


Db 508 CTGGGATCATACCTCTCGAGCTCTCAAGTGGAGAGTCTCCACGCGTGGCCATG 567
Qy 272 -----SerPheArgPheTyrPheValSerLeuAlaPhePheGlyPheLeuLeuPro 289
Db 568 TGGCGCGTGTCTCTTACCACTTCTCATCTGCTG-----TTCTCATCCCG 615
Qy 290 PheVallelelePheCysTyrThrThrLeuLeuHisLysLeu----- 304
Db 616 TTCTGTATCAGCTGGTGTGTACACGGCCACCATCTCAAGCTGTGGCGAGGAG 675
Qy 305 -----LysSerLysAspArgileTyrLeuGlyTyrleLysAlaValLeuLeu 321
Db 676 GCGACGCGGAGAGCGAGCGCGCGTGGCGCTGGCGCGGTGTCTGCTG--- 732
Qy 322 LeuVallePheThrleCysPheAlaProThrAsnleileLeuValleHisAla 341
Db 733 -----GCTTTGTACCTGCTTCCGCCCAACAACTTCTGCTCTCGGCGCATCTGT 786
Qy 342 AsnTyrTyrThrHisAsnThrAspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeu 361
Db 787 AGCGCGCTGTCTACGGC---AGAGGCTACTACCACTGTACAGCTCAAGCTGTGTCTC 843
Qy 362 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 844 AGTGCCTCAACAACTCTCTGGACCCGTTTGTATTACTTTCGTCCTCCCG 894
RESULT 11
US-09-943-718-3
; Sequence 3, Application US/09943718
; Patent No. US20020103361A1
; GENERAL INFORMATION:
; APPLICANT: Hufine, Constance F.
; Rossi, Devora L.
; Capone, Myriam
; Hedrick, Joseph A.
; Vicari, Alain
; Gorman, Daniel M.
; Zlotnik, Albert
; TITLE OF INVENTION: Mammalian Chemokines; Receptors;
; Reagents; Uses
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 30-Aug-2001
; APPLICATION NUMBER: US/09/943,718
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/009,817
; FILING DATE: 20-APR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0588K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2588 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1080
FEATURE:
NAME/KEY: misc_feature
LOCATION: 158
OTHER INFORMATION: /note= "residues 158, 159, and 276
probably absent, changing reading frame between those positions;
sequences provided in SEQ ID NO: 5 and 6"
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-943-718-3
Alignment Scores:
Pred. No.: 2,666-34 Length: 2588
Score: 471.50 Matches: 119
Percent Similarity: 52.30% Conservative: 74
Best Local Similarity: 32.25% Mismatches: 155
Query Match: 22.07% Indels: 23
DB: 10 Gaps: 6
US-09-208-629f-3 (1-407) x US-09-943-718-3 (1-2588)
Qy 22 LeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
Db 1040 CTGGAAGCGGAGGACTGCAATTTGCTCTCTACACTTCTGTGACTCAGC-ATCAGCGCTG 1098
Qy 42 SerAspAsnSerAlaLysProThrLeu-----ThrIleLysSerPheAsnGly----- 57
Db 1099 GCAGAGGGCATCCAGACCCCGACATCTACGATGATGATAAGAGTACCAGGGGAAGCCAT 1158
Qy 58 -----GlyProGlnAsnThrPheGluGluPheProLeuSerAsp----- 70
Db 1159 GAAGGCCCTCTGGTCCCACTACACTCAAGGACCGGAGCTCTCAGACAGCCTAAT 1218
Qy 71 IleGluGlyTyrThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIle 90
Db 1219 CCACGAGGCTACCCGGGCAAAITC-----TGTCACACGACAGT--- 1257
Qy 91 SerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGln 110
Db 1258 GACACGCTGAGCTCCCGCCAGCTCTCAAGCACTCTCTGGGGTGGGTCCCAAG 1317
Qy 111 ValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThr 130
Db 1318 CTGTACTCTGCTCTATGGGCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1377
Qy 131 LeuTyrPysLeuSerLeuArgThrLysSerIleSerLeuValIlePhe-HisThrAsnLe 150
Db 1378 CTGTGGGTGTGGCCACAAGGGTGCCACGCTGCCATCCACCATCTGTCTCATGGAACCT 1437
Qy 150 uAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAs 170
Db 1438 GGCAGTGGTGTATCTGTCTGTGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1497
Qy 170 nGlyAsnAsnTyrValPheGlyGluValMetCysArgIleThrThrValValPheTyrGl 190
Db 1498 TGCCACAGCGCTGCCATTTGGTGAGGCTGCTGCCGGGTGGCCACAGCTGCCCTCTATGG 1557
Qy 190 yAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaTh 210
Db 1558 CCACATGTATGGTTTCAGTGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1617
Qy 210 rAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGl 230
Db 1618 GGTGCATCTCTTTCGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1677
Qy 230 yIleValTyrValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTy 250
Db 1678 GGTGGCGCTGCTCTCTGCAGCCACCCCTGGCTTGGCTTCTCTACTCTCTGCAGCGG 1737
Qy 250 rHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerPr 270


```
QY 142 ---SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 190 AACTCTACCAACCCTCTATTCAACAAATTTGGTGATTTCTGATATCTATTTTACCACGGCT 249
QY 161 LeuProPheLysIleAlaIleHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 250 TTGCGTACACGAATAGCTACTACTGCAATGGCTTTGACTGGAGATCGGAGATCGCTTG 309
QY 181 CysArgIleThrValValPheThrGlyAsnMetTyrCysAlaIleLeuLeuLeuThr 200
Db 310 TGATAGGATAACTGGCTAGTGTGTTTATCATCAACATATCCAGGTGGAACCTTTATGACC 369
QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 370 TGCTGAGTATTGACCGCTTCATTTGCTGTGGTGCACCTCTACGCTACACAGATAA 429
QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 430 AGGATTGAACATAGCAAAAGCGTGCATATTGCTGGATTTCTAGTATTGCTCAGACA 489
QY 241 LeuProPheValIle-----LeuLysGlnGluTyrHisLeuValHisSerGluIle 257
Db 490 CTCCTCATCTCTCAACCTCTGCTCAAGCAGAG-----GCTGAAGG 534
QY 258 ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 535 ATTACATGCATGGAGTATCCAACTTTGAAGAACTAAATCT-----CTTCCCTGGATT 588
QY 278 PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
Db 589 CTGCTTGGGCGATGTTTTCATAGGATATGTTACTTCCACTTATAATCTCATCTCTGCTAT 648
QY 298 ThrThrLeuIleHisLys-----LeuLysSerLysAsp 308
Db 649 TCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACCAACCACTCACTGAGAATCT 708
QY 309 ArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCys 328
Db 709 GGTGTAAACAAAGGCTCTCAACAAATTTCTTATTATTATTGTTGTTGTTCTCTGT 768
QY 329 PheAlaProThrAsnIleLeuValIleHis-----HisAlaAsn 342
Db 769 TTCACACCTTACCATTGTTGCAATTTATTCACACATATGATTAAAGAGCTTCTCTCTAAT 828
QY 343 Tyr-----TyrTyrHisAsnThrAsp---SerLeuTyrPheMetTyrLeuIle 357
Db 829 TTCTCGAATGTAGCCAAAGACATCTGCTCCAGATTTCTCTGCACITTT-----876
QY 358 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSer 377
Db 877 ACAGTATGCTGATGAATTCATTTGCTGTCATGGACCTTTTATCTACTTCTTTGCAATGT 936
QY 378 Lys 378
Db 937 AAA 939

RESULT 15
US-10-225-567A-111
; Sequence 111, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmet, Glena C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: Patent in version 3.1
```

```
; SEQ ID NO 111
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-111

Alignment Scores: 7.59e-31 Length: 1638
Pred. No.: 431.50 Matches: 95
Score: 52.82% Conservative: 64
Percent Similarity: 31.56% Mismatches: 107
Best Local Similarity: 20.20% Indels: 35
Query Match: 9 Gaps: 9
DB:

US-09-208-629f-3 (1-407) x US-10-225-567A-111 (1-1638)

QY 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 121
Db 103 TATGCACATCACAGCAGCGCCAGGATAGTAATAGCTCTGCATTACAGCCCTCGTCTTCATC 162
QY 122 ValGlyValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLysSerIle 141
Db 163 ATTGGGCTCGTGGGAACTTACTAGCTTGGTGGTCAITTTGTTCAAAACAGGAAAAAATC 222
QY 142 ---SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 223 AACTCTACCAACCCTCTATTCAACAAATTTGGTGATTTCTGATATCTATTTTACCACCGCT 282
QY 161 LeuProPheLysIleAlaIleTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 283 TTGCTCTACAGTAATAGCTTACTATGCAATGGGCTTTGACTGGAGAAATCGGAGATGCTCTG 342
QY 181 CysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThr 200
Db 343 TGATAGTAATCTCGCTAGTGTGTTTACATCAACATATGCAAGTGTGAACCTTTATGACC 402
QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 403 TGCTGAGTATTGACCGCTTCATTTGCTGGTGCAACCTCTACGCTTACAACAGATAAAA 462
QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 463 AGGATTGAACATGCAAAAGCGCTGTGCATATTTGCTGGATTTCTAGTATTGCTCAGACA 522
QY 241 LeuProPheValIle-----LeuLysGlnGluTyrHisLeuValHisSerGluIle 257
Db 523 CTCCTCATCTCAACCTATGTCAAAGCAGGAG-----GCTGAAGG 567
QY 258 ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 568 ATTACATGCATGGAGTATCCAACTTTGAAGAACTAAATCT-----CTTCCCTGGATT 621
QY 278 PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
Db 622 CTGCTTGGGCGATGTTTTCATAGGATATGTTACTTCCACTTATAATCTATCTCTCTCTAT 681
QY 298 ThrThrLeuIleHisLys-----LeuLysSerLysAsp 308
Db 682 TCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACCAACCACTCACTGAGAATCT 741
QY 309 ArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCys 328
Db 742 GGTGTAAACAAAGGCTCTCAACAAATTTCTTATTATTATTGTTGTTGTTCTCTGT 801
QY 329 PheAlaProThrAsnIleIleLeuValIleHis-----HisAlaAsn 342
Db 802 TTCACACCTTACCATTGTTGCAATTTATTCACATATGATTAAAGAGCTTCTGTTTCTCTAAT 861
QY 343 Tyr-----TyrTyrHisAsnThrAsp---SerLeuTyrPheMetTyrLeuIle 357
Db 862 TTCCTGGAATGTAGCCAAAGACATTCGTTCCAGATTTCTCTGCACITTT-----909
QY 358 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSer 377
```

```

Db 910 ACAGTAGCCTGATGAACCTCAATTGCTGATGAGACCTTTATCTACTCTTTGCGATG 969
    |||
Qy 378 Lys 378
    |||
Db 970 AAA 972

RESULT 16
US-10-251-385-205
; Sequence 205, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; PRIOR FILING DATE: 2002-09-20
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 205
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-205

Alignment Scores:
Pred. No.: 8e-31 Length: 1086
Score: 428.50 Matches: 96
Percent Similarity: 51.78% Conservative: 64
Best Local Similarity: 31.07% Mismatches: 98
Query Match: 20.06% Indels: 51
DB: 9 Gaps: 10

US-09-208-629f-3 (1-407) x US-10-251-385-205 (1-1086)

Qy 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 121
Db 70 TATGACATCACAGACGGCGAGGATAGTAATGCTCTGCATTACAGCCCTCGTCTTCATC 129
Qy 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerIle 141
Db 130 ATTGGGCTGTGGAAACTTACTAGCCTTGCTGCTCAATTGTTCAAAACAGGAAAAAATC 189
Qy 142 ---SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 190 AACTCTACCACCTCTATTCAACAAATTTGGTGATTTCTGATATCTTTTACCACGGCT 249
Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 250 TTGCTACACGAATAGCCTACTATGCAATGGGCTTTGACTGGGAATCGGAGATCGCTTG 309
Qy 181 CysArgIleThrThrValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThr 200
Db 310 TGTAGGATACTGCGCTAGTGTATTACATCAACATATGAGGTGGAATTTATGACC 369
Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 370 TGCCTGAGTATTGACCGCTTCATTGCTGTGTGTCACCTCTACGCTACACAGATAAAA 429
Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 430 AGGATTGAACATGCAAAAGGCGTGTGCATATTTGCTCTGGATTTCTAGTATTGCTCAGACA 489
Qy 241 LeuProPheValIle-----LeuLysGlnGlnTyrHisLeuValHisSerGluIle 257
Db 490 CTCCCCTCTCATCAACCCCTATGTCAAAGCAGGAG-----GCTGAAAGG 534
```

```

Qy 258 ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
    |||
Db 535 ATTACATGATGGAGTATCCAAACTTTGAAGAACTTAAATCT-----CTTCCCTGGATT 588
    |||
Qy 278 PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
    |||
Db 589 CTGCTTGGGGCATGTTTCATAGATATGTAATCTTCCACTTATATCAATCTCTCTCTCTAT 648
    |||
Qy 298 ThrThrLeuIleHisLysLeu----- 304
    |||
Db 649 TCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAACCCACTCACTGAGAAATCT 708
    |||
Qy 305 -----LysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 320
    |||
Db 709 GGTGTAAACAAAAGCTTAAAC-----ACAATTAATCTT 744
    |||
Qy 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHis--- 339
    |||
Db 745 ATTATTGTTGTTGTTCTCTCTGTTTACACCTTACCATGTTGCAATTAATTCACATATG 804
    |||
Qy 340 -----HisAlaAsnTyr-----TyrTyrHisAsnThrAsp--- 349
    |||
Db 805 ATTAAGAAGCTTCGTTTCTCTAATTTCTGGAATGTAGCCAAAGACATTCGTTCCAGATT 864
    |||
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
    |||
Db 865 TCTCTGCACCTT-----ACAGTATGCTGATGAACCTTCAATTCATGCTCATGGAC 912
    |||
Qy 370 ProPheLeuTyrPheValMetSerLys 378
    |||
Db 913 CCTTTTATCTACTCTTTTGCATGTAAA 939

RESULT 17
US-10-225-567A-303
; Sequence 303, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 303
; LENGTH: 2070
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-303

Alignment Scores:
Pred. No.: 1.3e-27 Length: 2070
Score: 398.50 Matches: 99
Percent Similarity: 45.43% Conservative: 60
Best Local Similarity: 28.29% Mismatches: 136
Query Match: 18.66% Indels: 55
DB: 9 Gaps: 9

US-09-208-629f-3 (1-407) x US-10-225-567A-303 (1-2070)

Qy 42 SerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGly-----Gly 58
    |||
Db 141 TCTGACTCCAGC-----CAAGCATGAATGCGCTTGAATGGCT 179
    |||
Qy 59 ProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThr 78
    |||
Db 180 CCCCCAGGCTGATCATCAACAACTTCTCCCTGGCCACGGCAGAG----- 221
```

```
QY 79 ThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAla 98
Db 222 -----CAATGTGGCCAGGAG----- 236
QY 99 ThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeu 118
Db 237 -----ACGCCACTGGAGAACATGCTGTTCGCCTCTCTACCTTCTG 278
QY 119 LeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr 138
Db 279 GAITTTATCTGGCTTTAGTTGGCAATACCTCTGTGGCTTTTCATCCGAGACCAC 338
QY 139 LysSer---IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe 157
Db 339 AAGTCCGGGACCCGGCCAAAGTGTCTCTGATGATCTGGCCGCTGTGCTGC 398
QY 158 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnTrpValPheGly 177
Db 399 GTGCTGTCTGCCACCCGCTGGTCTACCACTTCTCTGGGAACCACTGGCCATTGGG 458
QY 178 GluValMetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu 197
Db 459 GAAATCCATGCCGTCTACCGGCTTCTCTTACCTCAACATGACGCCGCACTTAC 518
QY 198 IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln 217
Db 519 TTCCTACCTGCATCAGCGCCGACCGTTTCTGGCCATTGTGCACCGGTCAAGTCCCTC 578
QY 218 LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPhe 237
Db 579 AAGCTCCGAGGCCCTCTACGCACACCTGGCTGTGCTCTCTGTGGTGGTGGTGGCT 638
QY 238 LeuTyrMetLeuProPheValIleLeuLysGlnIleTyrHisLeuValHisSerGluIle 257
Db 639 GTGGCCATGGCCCGCTGTGGTGAGCCACAG-----ACCGTG 677
QY 258 ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr 277
Db 678 CAGACCAACACACAGCGGTGGTCTGCCTGCAGCTGTACCGGAGAGGCTCCACCATGCC 737
QY 278 PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr 297
Db 738 CTGGTGTCCCTGGCA---GTGGCCTTCCACTTCCCGTTTCATCACCAGGTCACTGTCTAC 794
QY 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
Db 795 CTGCTGATATCCGACCGCTGGCGCAGGCTGTGTGGAGAGCGCTCAAGACCAAG 854
QY 314 TyrIleLysAlaValLeuLeuValIlePheThrIleCysPheAlaProThrAsn 333
Db 855 GCAGTGGCATGATCGGCATAGTGTGGCCATCTTCTGTGCTGTCTGTGGCCCTACCAC 914
QY 334 IleIleLeuValIleHisHisAlaAsnTyrTyrHisAsnThrAspSer----- 350
Db 915 GTCAACCGCTCCGCTACGTGCTGCTACCTACCGCAGCATGGGCTCTCTCGCCACCCAG 974
QY 351 -----LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 975 CGCATCTGGCCCTGGCAACCGCATCACCTCTGCTTCCAGCTCAACGGGGCACTC 1034
QY 369 AspProPheLeuTyrPheValMetSerLys 378
Db 1035 GACCCCATCATGATTCTTCTGTGGCTGAG 1064
```

RESULT 18

```
US-10-251-385-31
; Sequence 31, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
```

```
; TITLE OF INVENTION: Protein-Coupled
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-31

Alignment Scores:
Pred. No.: 9,95e-28 Length: 1020
Score: 395.00 Matches: 88
Percent Similarity: 50.00% Conservative: 53
Best Local Similarity: 31.21% Mismatches: 123
Query Match: 18.49% Indels: 18
DB: 5 Gaps: 5
```

US-09-208-629f-3 (1-407) x US-10-251-385-31 (1-1020)

```
QY 107 LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer 126
Db 85 CTGGAGAACATGCTGTTCGCCCTCTTCTACCTTCTGGATTATCTCTGGCTTTAGTTGGC 144
QY 127 AsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSer---IleSerLeuValIle 145
Db 145 AATACCTCGCTCTGTGGCTTTTATCCGAGACCAACAGTCCGGGACCCCGGCAACGTG 204
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
Db 205 TTCTCTGATCATCTGGCCGTGGCCGACTTGTCTGCTGCTGTCTGTCTGTCTGTCTGTCTG 264
QY 166 AlaTyrHisLeuAsnGlyAsnTrpValPheGlyGluValMetCysArgIleThrThr 185
Db 265 GTCTACCATCTTCTGGGAACCACTGGCCATTGGGGAATTCGATCGCTCTACCGGC 324
QY 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsn 205
Db 325 TTCTCTCTTACCTCAACATGTAGCCAGCATCTACTTCTCACCTGCATCAGCGCCGAC 384
QY 206 ArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSer 225
Db 385 CGTTTCTGTGGCCATTGTGCACCGGTCAAGTCCCTCAAGTCCGAGGCCCCCTCTACGCA 444
QY 226 LeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIle 245
Db 445 CACCTGGCTGTGCCCTTCTGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 504
QY 246 LeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAsp 265
Db 505 AGCCACAG-----ACCGTCAGACCAACACACACCGGTGTGTCTGC 543
QY 266 AlaCysGluSerProSerSerPheArgPheTyrPheValSerLeuAlaPhePheGly 285
Db 544 CTGACGCTGTACCGGAGAGGCTCCACCATGCTGTGTCTGTCTGTCTGTCTGTCTGTCTG 600
QY 286 PheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHisLysLeuLys 305
Db 601 TTCACCTTCCGTTTCATACCACCGGTCACTGCTACTCTGTCTGTCTGTCTGTCTGTCTGT 660
QY 306 Ser-----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIle 321
Db 661 CAGGGCTCGTGTGGAGAGCGCTCAAGACCAAGGAGTGGGATGATGATGATGATGATGATG 720
QY 322 LeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHisAla 341
Db 721 CTGGCCATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
```



```
Db 205 TTCTGATGATCTGGCGGCGCGCGCTTGTGCGTGTCTGCGCCACCCGCGT 264
      ::::::::::::::::::::
Qy 166 AlaTyrHisLeuAenGlyAenAsnTrpValPheGlyGluValMetCysArgIleThr 185
      ::::::::::::::::::::
Db 265 GTCTACCACTTCTTGGGAACCACTGGCCATTTGGGAATTCGATCGCGTCTCACCGGC 324
      ::::::::::::::::::::
Qy 186 ValValPheTyrGlyAenMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsn 205
      ::::::::::::::::::::
Db 325 TTCTCTTCTACCTCAACATGACCGCAGCATCTACTTCTCCTACCTGCACTACCGCGGAC 384
      ::::::::::::::::::::
Qy 206 ArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSer 225
      ::::::::::::::::::::
Db 385 CGTTCTTCGCCATTTGGCACCCTCGTCAAGTCCCTCAAGTCCGCGAGGCCCTCTACGCA 444
      ::::::::::::::::::::
Qy 226 LeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIle 245
      ::::::::::::::::::::
Db 445 CACCTGGCCCTGTGCTTCTTCTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 504
      ::::::::::::::::::::
Qy 246 LeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAsp 265
      ::::::::::::::::::::
Db 505 AGCCACAG-----ACGTGCAGACCAACACACCGTGGTGTGC 543
      ::::::::::::::::::::
Qy 266 AlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGly 285
      ::::::::::::::::::::
Db 544 CTGCAGCTGTACCGGAGAGGCTCCACCACCTGCTGTGTGTGTGTGTGTGTGTGTGT 600
      ::::::::::::::::::::
Qy 286 PheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHisLysLeuLys 305
      ::::::::::::::::::::
Db 601 TTCACCTTCCGCTTCATCACCAAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
      ::::::::::::::::::::
Qy 306 Ser-----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIle 321
      ::::::::::::::::::::
Db 661 CAGGCGCTGGTGGAGAGCGCTCAAGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGTGT 720
      ::::::::::::::::::::
Qy 322 LeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHisAla 341
      ::::::::::::::::::::
Db 721 CTGGCCATCTCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
      ::::::::::::::::::::
Qy 342 AsnTyrTyrTyrHisAsnThrAspSer-----LeuTyrPheMetTyrLeu 356
      ::::::::::::::::::::
Db 781 CACTACCGCAGCCTGCGGCTCTCTGCGCCACCCAGCGCATCTCTGGCCCTGGCAACCCG 840
      ::::::::::::::::::::
Qy 357 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMet 376
      ::::::::::::::::::::
Db 841 ATCACCCTTCCCTCACCGCCTCAACGGGCGCTACCGGCGCTACCGCCCATCATGTATTTCT 900
      ::::::::::::::::::::
Qy 377 SerLys 378
      ::::::::::::::::::::
Db 901 GCTGAG 906
      ::::::::::::::::::::
```

RESULT 21

```
US-10-225-567A-224
; Sequence 224, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225.567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 224
; LENGTH: 2299
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-225-567A-224
```

```
Alignment Scores:
Pred. No.: 1,318-26 Length: 2299
Score: 388.50 Matches: 102
Percent Similarity: 47.15% Conservative: 72
Best Local Similarity: 27.64% Mismatches: 126
Query Match: 18.19% Indels: 69
DB: 9 Gaps: 14
```

US-09-208-629F-3 (1-407) x US-10-225-567A-224 (1-2299)

```
Qy 35 CysGlnSerGlyIleAsnValSerAspAsnSerAlaLysProThrLeuThrIleLysSer 54
      ::::::::::::::::::::
Db 16 TGTGAGAGTGGT-----GAACCCCTGCAGCCAGCAGGCGCTCTCGAAAAA 60
      ::::::::::::::::::::
Qy 55 PheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGlyTrp 74
      ::::::::::::::::::::
Db 61 AAGTCCATGGGTGACAGAAGA---TTCATTGATTTCCCAATTC----- 99
      ::::::::::::::::::::
Qy 75 ThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu--- 93
      ::::::::::::::::::::
Db 100 -----CAAGATTCAAATTCAGGCTCAGA 123
      ::::::::::::::::::::
Qy 94 ---HisValAsnAsnAlaThrIleGly-----TyrLeuArgSerSerLeuSerThr 109
      ::::::::::::::::::::
Db 124 CCAGGTTGGCAATGCTACTGCCAATAATACTTGCATTGTTGATGATCTCTCAAGTAT 183
      ::::::::::::::::::::
Qy 110 GlnValIleProAlaIleTyrIleLeuPheValGlyValProSerAsnIleVal 129
      ::::::::::::::::::::
Db 184 AATCTCAATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 243
      ::::::::::::::::::::
Qy 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSer---LeuValIlePheHisThr 148
      ::::::::::::::::::::
Db 244 TCTCTGTTTGTCTTCTGTTTCCGATGAAATAGAGAGTACGACTGCTATTTTATCACC 303
      ::::::::::::::::::::
Qy 149 AsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHis 168
      ::::::::::::::::::::
Db 304 AATCTAGCTGTCTGATTGTTGTTTGTCTGTACACTACCTTTTAAATAATTTTACAAC 363
      ::::::::::::::::::::
Qy 169 LeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPhe 188
      ::::::::::::::::::::
Db 364 TTCAAC---CGCCACTGGCGCTTTTGGTGACACCCCTCTGCAAGATCTCTGCAACTG 420
      ::::::::::::::::::::
Qy 189 TyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeu 208
      ::::::::::::::::::::
Db 421 CTTACCAACATCTATGGGAGCATGCTCTTCTCACCTGTATAGTGTGGATCGTTTCTTG 480
      ::::::::::::::::::::
Qy 209 AlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMet 228
      ::::::::::::::::::::
Db 481 GCCATTGTCTATCTTTTCGATCTCGTACTATTAGGACTAGGAGGAATCTGCGCATTTG 540
      ::::::::::::::::::::
Qy 229 CysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGln 248
      ::::::::::::::::::::
Db 541 TGTGCTGGTGTCTGGATCTCTAGTCCCTAGTGGCGGTATTTTCAGCTCTTTGTTT--- 594
      ::::::::::::::::::::
Qy 249 GluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGlu 268
      ::::::::::::::::::::
Db 595 -----TCCACCACCTAATGTCAACAATGCAACCAACCACCTCTCTT 633
      ::::::::::::::::::::
Qy 269 SerProSerSerPheArgPheTyr-----TyrPheValSerLeuAlaPhePhe--- 284
      ::::::::::::::::::::
Db 634 GAAGGCTCTTCCAAACGCTGTCTGGAAGACTTTATTTATCCAAGATCAACAATTTATTGAA 693
      ::::::::::::::::::::
Qy 285 -----GlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHis 302
      ::::::::::::::::::::
Db 694 GTTGTGGGTTTATCATCTCTTAATATTGAATGTCTTGTCTTCTTGTGGTGGCTGAGA 753
      ::::::::::::::::::::
Qy 303 LysLeuLys-----SerLysAspArgIleTrpLeu 312
      ::::::::::::::::::::
Db 754 ACTCTTGGCAAGCCTCTACTCTGCTCAAAATGGGACCAATGAAGAAAAAGTA----- 807
      ::::::::::::::::::::
Qy 313 GlyTyrIleLysAlaValLeuLeuIlePheThrIleCysPheAlaProThr 332
      ::::::::::::::::::::
```


Db 808 -----CTGAAATGATCAGTACATATGCGAGTCTTTGGTATGCTTTGTACCCCTAC 861
Qy 333 AsnIleIleLeuValIleHisHis-----AlaAsnTyrTyr 345
Db 862 AACTGCTCTCTCTTGTATGCGCTGGTCCCAAGCTATTACTAATGCTTTTG 921
Qy 346 HisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 365
Db 922 GAAAGATTTTCAAGATC-----ATGTACCAATCACCTTGTGCTTGCCTGCAACTCTGAAC 975
Qy 366 SerCysLeuAspPheProPheLeuTyrPhe 374
Db 976 TGTGTTTGGACCTTTCATCTATTAC 1002

RESULT 22

US-09-826-791-1
; Sequence 1, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1el Polypeptide
; FILE REFERENCE: PC10914ADAM
; CURRENT APPLICATION NUMBER: US/09/826,791
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 0008504.3
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/198,367
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-826-791-1

Alignment Scores:
Pred. No.: 3,35e-26 Length: 993
Score: 378.50 Matches: 88
Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: 10 Gaps: 8

US-09-208-629F-3 (1-407) x US-09-826-791-1 (1-993)

Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 7 CCAATGGCACCTTCAGCAATAACACAGCAGCAACTGCACAATT-----GAA 54
Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Db 55 AACTTCAAGAGAGAAATTTTCCCAATGTATATCTGATATATTTTCTGGGAGTCTTG 114
Qy 126 SerAsnIleValThrLeuTyrLys----LeuSerLeuArgThrLysSerIleSerLeuVal 144
Db 115 GGAATGGGTGTCCATATATGTTTCTCGCAGCTTATAAAGAAGTCCACATCTGTGAAC 174
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 175 GTTTTCATGTAATATGGCCATTTTCAGATCTCTGTTCATAGCAGCTTCCTCCCTCAGG 234
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCysAsnGlyThr 184
Db 235 GCTGACTATTATCTTAGAGCTCAATGTATATTTGGAGACCTGGCTCGAGATATG 294
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 295 TCTATTCTCTGTATGTCAACATGTACAGCAGTATTATTCTCGACCGGTGCTGAGTGT 354
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 355 GTGCGTTTCTGGCAATGGTTTCCCTTTCACCCCTTCTCGCATGTCCACCATCAGGAGT 414

Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 415 GCGTGGATCCTCTGTGGGATCATATGGATCCTTATC-----ATGGCTTCTCAATA 465
Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 466 ATGCTCTCGGACAGTGGCTCTGAGCAGAACGCGCAGTGTACATCATGCTTAGAGCTG--- 522
Qy 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
Db 523 -----AATCTTATAAATTGCTAAGCTGAGCAGCATG 555
Qy 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
Db 556 AACTATATTGCTTGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 615
Qy 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleThrLeuGly 313
Db 616 CTGCTGATCATTCGGGTTCTGTTAAAGTGGAGTCCAGAAATCGGGGCTCGGGGTTTCT 675
Qy 314 TyrIleLysAla-----ValLeuLeuIleValIlePheThrIleCysPheAla 330
Db 676 CACAGGAGGCACTGACCACTATCATCATCATCTTGTATCTTCTTGTGTTTCTG 735
Qy 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrHis----AsnThrAsp 349
Db 736 CCCTATCACACACTGAGGACCGTCCACTTGCAGCATGGAAGTGGGTTTATGCAAGAC 795
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 796 AGACTGCATAAAGCTTTGTTTATCATCACTGGCTTGGCAGCAGCAATGCTGCTTCAAT 855
Qy 370 ProPheLeuTyrPhe 374
Db 856 CCTGCTCTATTAC 870

RESULT 23

US-09-828-478-1
; Sequence 1, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-828-478-1

Alignment Scores:
Pred. No.: 3,59e-26 Length: 1041
Score: 378.50 Matches: 88
Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: 9 Gaps: 8

US-09-208-629F-3 (1-407) x US-09-828-478-1 (1-1041)

Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 55 CCAATGGCACCTTCAGCAATAACACAGCAGCAACTGCACAATT-----GAA 102

```
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 103 AACTTCAGAGAGAAATTTTCCCAATGTATATCTGATATATTTTCTGGGAGTCTTG 162
QY 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
DB 163 GGAAATGGGTGTCCATATATGTTTCTCGAGCTTATAAGAAGTCCACATCTGTGAAC 222
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
DB 223 GTTTTCATGCTAAATCTGGCCATTTCCAGATCTCTGTTCATAGACGCTTCCCTTCAGG 282
QY 165 IleAlaTyrHisLeuAsnGlyValPheGlyGluValMetCysArgIleThr 184
DB 283 CGTGCATATTATCTTAGAGCTCCCAATGGATATTGGAGACCTGGCCTGCAGATTATG 342
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
DB 343 TCTTATTCCTGTGATGCTCAACATGTACACAGTATTATTCTCGACCGTTCCTTCAGG 402
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 403 GTGCGTTTCTGGCAATGGTTCACCCCTTTCGGCTTCTGCATGTCCACGACATCAGGAGT 462
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 463 GCCTGCATCTCTGTGGGATCATATGATCCTTATC-----ATGCTTCTCTCAATA 513
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
DB 514 ATGCTCTCGACAGTGGCTCTCGAGAGACGGCAGTGTCCATCATGCTTAGAGCTG--- 570
QY 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
DB 571 -----AATCTCTATAAAATTCCTAAGCTGCAGACCATG 603
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr 297
DB 604 AACTATATTGCTTGTGTGGCTGCTGCTGCATTTTTCACATCAGCATCTGTTAT 663
QY 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
DB 664 CTGCTGATCATCTGGGTCTCTGTTAAAGTGGAGTCCCAAGATCGGGCTGCGGTTCCT 723
QY 314 TyrIleLysAla-----ValLeuLeuIleValIlePheThrIleCysPheAla 330
DB 724 CACAGGAAGGCATCTGACCACCATCATCATCCTTGATCATCTTCTTGTGTTTCTCTG 783
QY 331 ProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
DB 784 CCCTATCACACACTGAGACCGCTCCATCTGACGACATGGAAGTGGTTTATGCAAGAC 843
QY 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
DB 844 AGACTGCATAAAGCTTTGGTTTATCACACTGSCCTTGGCAGCAGCAATGCTTCTTCAAT 903
QY 370 ProPheLeuTyrPhe 374
DB 904 CCTCTGCTCTATTAC 918
```

RESULT 24

```
US-09-826-791-5
; Sequence 5, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1el Polypeptide
; FILE REFERENCE: P010914ADAM
; CURRENT APPLICATION NUMBER: US/09/826,791
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 0008504.3
; PRIOR FILING DATE: 2000-04-05
```

```
; PRIOR APPLICATION NUMBER: 60/198,367
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
```

```
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1041)
US-09-826-791-5
```

```
Alignment Scores: 3 59e-26 Length: 1041
Pred. No.: 378-50 Matches: 88
Score: 378-50 Conserv: 63
Percent Similarity: 49.51% Mismatches: 121
Best Local Similarity: 28.85% Indels: 33
Query Match: 17.72% Gaps: 8
DB: 10
```

```
US-09-208-629F-3 (1-407) x US-09-826-791-5 (1-1041)
```

```
QY 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
DB 55 CCAATATGGCACCTTCAGCAATAAACACAGCAGCAACTGCACAATT-----GAA 102
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 103 AACTTCAAGAGAGATTTTCCCAATGTATATCTGATAAATATTTTCTGGGAGTCTTG 162
QY 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
DB 163 GGAAATGGGTGTGTCATATATGTTTCTGAGACCTTATAAGAAGTCCACATCTGTGAAC 222
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
DB 223 GTTTTCATGCTAAATCTGGCCATTTCCAGATCTCTGTTTTCATAGCAGCGTTCCTTCAGG 282
QY 165 IleAlaTyrHisLeuAsnGlyValPheGlyGluValMetCysArgIleThr 184
DB 283 GGTGACTATTATCTTAGAGCTCCCAATGGATATTGGAGACCTGGCCTGCAGGATTATG 342
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
DB 343 TCTTATTCCTGTGATGCTCAACATGTACACAGTATTATTCTCGACCGTTCGTGAGTGT 402
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 403 GTGCGTTTCTGGCAATGGTTCACCCCTTTCGGCTTCTGCATGTCCACGACATCAGGAGT 462
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 463 GCCTGCATCTCTGTGGGATCATATGATCCTTATC-----ATGCTTCTCTCAATA 513
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
DB 514 ATGCTCTCGACAGTGGCTCTCGAGAGACGGCAGTGTCCATCATGCTTAGAGCTG--- 570
QY 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
DB 571 -----AATCTCTATAAAATTCCTAAGCTGCAGACCATG 603
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr 297
DB 604 AACTATATTGCTTGTGTGGCTGCTGCTGCATTTTTCACATCAGCATCTGTTAT 663
QY 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
DB 664 CTGCTGATCATCTGGGTCTCTGTTAAAGTGGAGTCCCAAGATCGGGCTGCGGTTCCT 723
QY 314 TyrIleLysAla-----ValLeuLeuIleValIlePheThrIleCysPheAla 330
DB 724 CACAGGAAGGCATCTGACCACCATCATCATCCTTGATCATCTTCTTGTGTTTCTCTG 783
```

Db 724 CACAGGAGGCACTGACCACCATCATCATCACCTTGATCATCTTCTTCTGTTCTCCTG 783
Qy 331 ProThrAsnIleLeuValIleHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
Db 784 CCCTATCACACACTGAGGACCGTCCACTTCAGCACATGGAAGTGGGTTTATGCAAGAC 843
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 844 AGACTGCATAAAGCTTTGGTTATACACATGGCTTGGCAGCAGCAATGCTGCTCAAT 903
Qy 370 ProPheLeuTyrPhe 374
Db 904 CCTCTGCTCTATTAC 918
RESULT 25
US-09-866-230-6
; Sequence 6, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US/09/866,230
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1041)
US-09-866-230-6
Alignment Scores:
Pred. No.: 3,59e+26 Length: 1041
Score: 378.50 Matches: 88
Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: 10 Gaps: 8
US-09-208-629f-3 (1-407) x US-09-866-230-6 (1-1041)
Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 55 CCAATGGCCACTTCAGCAATAACACAGCAGGAAGTGCACAATT-----GAA 102
Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Db 103 AACTTCAGAGAGAAATTTTCCCAATGTATATCTGATATATTTTCTGGGAGCTTG 162
Qy 126 SerAsnIleValThrLeuTyrLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
Db 163 GGAATGGGTTGTCATATGTTTCTCAGCCTTTATAAGAGTCCACATCTGTGAAC 222
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 223 GTTTTCATGTAATCGGCAATTCAGATCTCCTGTTCATAAGCAGCGCTTCCCTCAGG 282
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCysArgIleThr 184
Db 283 GCTGACTATTATCTTAGAGGCTCCAATGGATATTGGAGACTGGCCTCAGAGATTATG 342
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 343 TCTATTCTCTGATGTCAACATGTACAGCAGATTTATTCTCGACCGTCTGAGTGT 402
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB: |||||

Db 403 GTGCGTTTCTGTCGAATGGTTCCACCCCTTTGGCTTCTGCATGTCCACGAGCATCAGAGT 462
Qy 225 SerLeuLeuMetCysGlyIleValTyrPheValMetValPheLeuTyrMetLeuProPheVal 244
Db 463 GCCTGGATCCTCTGTGGGATCATATGATCCTTATC-----ATGGCTTCTCAATA 513
Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
Db 514 ATGCTCTGAGCAGTGGCTCTGAGCAGAACGGCAGTGTACATCATGTCTAGAGCTG-- 570
Qy 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
Db 571 -----AATCTCTATAAAATTCCTAAGCTGCAGACCATG 603
Qy 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
Db 604 AACTATATTGCGCTGGTGGCTGCTGCTGCCATTTTTCACACTCAGCATCTGTTAT 663
Qy 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTyrPheLeuGly 313
Db 664 CTGCTGATCATCTGGGTTCTGTTAAAGTGGAGGTCCAGAAATCGGGGCTGCGGTTTCT 723
Qy 314 TyrIleLysAla-----ValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
Db 724 CACAGGAGGCACTGACACCATCATCATCATCTTCTTGTGTTTCTCCTG 783
Qy 331 ProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
Db 784 CCTATCACACACTGAGGACCGTCCACTTCAGCAGATGGAAGTGGGTTTATCAGAAC 843
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 844 AGACTGCATAAAGCTTTGGTTATACACACTGGCTTGGCAGCAGCAATGCTGCTCAAT 903
Qy 370 ProPheLeuTyrPhe 374
Db 904 CCTCTGCTCTATTAC 918
RESULT 26
US-09-779-679-1
; Sequence 1, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sa
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26

```
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (105)..(1142)
; NAME/KEY: misc feature
; LOCATION: (120)..(122)
; OTHER INFORMATION: Alternative ATG start site
US-09-779-679-1

Alignment Scores:
Pred. No.: 4.73e-26 Length: 1260
Score: 378.50 Matches: 88
Percent Similarity: 49.51%
Best Local Similarity: 28.85%
Query Match: 17.72%
DB: 9
Gaps: 8

US-09-208-629F-3 (1-407) x US-09-779-679-1 (1-1260)

QY 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 159 CCAATGGCCACCTTCAGCAATAACAACAGCAGCACTGCACATT-----GAA 206
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Db 207 AACTTCAAGAGAGAAATTTTCCCAATTTGATATCTGATAATATTTTCTGGGAGTCTTG 266
QY 126 SerAsnIleValThrLeuTyrLys---LeuSerLeuArgThrIlySerIleSerLeuVal 144
Db 267 GGAATGGGTGTCCATATATGTTTCTCGACCTTATAAGAGTCCACATCTGTGAAC 326
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 327 GTTTCATGCTAAATCTGGCCATTCTCAGATCTCTGTTCATAAGCAGCTTCCTTCAGG 386
QY 165 IleAlaTyrHisLeuAenGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 387 GCTGACTATTATCTTAGAGGCTCCAATTTGGATATTTGGAGACCTGGCCTGCAGGATTATG 446
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 447 TCTTATTCTCTGTATGTCAACATGTACAGCAGTATTTATTTCTCGACCGTGTGAGTGT 506
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 507 GTGGTTCCTGGCAATGGTTCACCCCTTTCCGGCTTCTGCATGTCCACCATCAGGAGT 566
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 567 GCCTGGATCTCTCTGGGATCATATGATCTTATC-----ATGGCTTCTCTCAATA 617
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 618 ATGCTCTCGACAGTGGCTCTGACGAGACGGAGTGTCTACATCATCTTAGAGCTG--- 674
QY 265 AspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeu----- 281
Db 675 -----AATCTCTATAAATTCGTAAGCTGCAGACCATG 707
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
Db 708 AACTATATTGCTTGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
```

```
QY 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
Db 768 CTGCTGATCATTCGGGTCTGTAAAGTGGAGTCCCAAGATCGGGTCTCGGGTTCT 827
QY 314 TyrIleLysAla-----ValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
Db 828 CACAGGAGGCACTGACCACTCATCATCATCTTCTCTTCTTCTTCTTCTTCTTCT 887
QY 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
Db 888 CCTATCACACACTGAGGACCTGACATGTCACATGGAAGATGGTGTATATGCAAGAC 947
QY 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 948 AGACTGCATAAAGCTTTGGTTATACACTGGCTTGGCAGCAGCCATGCTGCTTCAAT 1007
QY 370 PropheLeuTyrPhe 374
Db 1008 CCTCTGCTCTATTAC 1022

RESULT 27
US-09-779-679-24
; Sequence 24, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sa
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-779-679-24

Alignment Scores:
Pred. No.: 4.73e-26 Length: 1260
Score: 378.50 Matches: 88
```

Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: Gaps: 8

US-09-208-629F-3 (1-407) x US-09-779-679-24 (1-1260)

QY 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
DB 159 CCAATGGCCTTACGCAATAACACAGCAGCAACTGCACAATT-----GAA 206
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 207 AACTTCAAGAGAGAAATTTTCCCAATGTATATCTGATAATATTTTCTGGGAGTCTTG 266
QY 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
DB 267 GGAATGGGTGTCATATATGTTTCTGCAGCTTATAAGAGTCCACATCTGTGAAC 326
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
DB 327 GTTTTCATGCTAAATCTGGCAATTCAGATCTCTGTTCATAAGCAGCTTCCCTTCAGG 386
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 194
DB 387 GCTGACTATATCTTAGAGGCTCCAATGTGATATTTGGAGACCTGGCCTGCAGGATATG 446
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
DB 447 TCTATCTCTGTGATGTCAACATGTACAGCAGATATTTATTTCTGACCGTGTGAGTGT 506
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 507 GTGCGTTTCTGCAATGGTTCACCTTTTCGGCTTCTGCATGTCACCATCAGGAGT 566
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 567 GCTGATCTCTGTGGGATCATATGATCTTATC-----ATGGCTTCTCCTCAATA 617
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
DB 618 ATGCTCTGCAGCAGTGGCTCTGACGACAGCGGAGTGTACATCATCTGTAGAGTGT--- 674
QY 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
DB 675 -----AATCTCTATAAAATTCCTAAGCTGCAGACCATG 707
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr 297
DB 708 AACTATATTGCTTGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
QY 298 ThrThrIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
DB 768 CTGCTGATCATCGGGTTCCTGTTAAAGTGGAGGTCCAGAAATGGGGCTGGGGTTCCT 827
QY 314 TyrIleLysAla-----ValLeuIleLeuValIlePheThrIleCysPheAla 330
DB 828 CACAGGAAGGCACTGACCAACATCATCACCTTGATCATCTTCTTGTGTGTTTCTGCT 887
QY 331 ProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
DB 888 CCCTATCACACACTGAGGACCGCTCCACTTGACGACATGGAAAGTGGGTTCATCAAGAC 947
QY 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
DB 948 AGACTGATAAAGCTTTGGTTTATACACTGCTGCTTGGCAGCACCATGCTGCTCAAT 1007
QY 370 ProPheLeuTyrPhe 374
DB 1008 CCTCTGCTATTATAC 1022

RESULT 28
US-09-828-478-3

; Sequence 3, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1430
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-828-478-3

Alignment Scores:
Pred. No.: 5,67e-26 Length: 1430
Score: 378.50 Matches: 88
Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: Gaps: 8

US-09-208-629F-3 (1-407) x US-09-828-478-3 (1-1430)

QY 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
DB 307 CCAATGGCCTTACGCAATAACACAGCAGCAACTGCACAATT-----GAA 354
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 355 AACTTCAAGAGAGAAATTTTCCCAATGTATATCTGATAATATTTTCTGGGAGTCTTG 414
QY 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
DB 415 GGAATGGGTGTCATATATGTTTCTGCGACCTTATAAGAGTCCACATCTGTGAAC 474
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
DB 475 GTTTTCATGCTAAATCTGGCAATTCAGATCTCTGTTTCATAAGCAGCTTCCCTTCAGG 534
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
DB 535 GCTGACTATATCTTAGAGGCTCCAATGTGATATTTGGAGACCTGGCCTGCAGGATATG 594
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
DB 595 TCTATTCCTGTGATGTCAACATGTACAGCAGATATTTATTTCTGACCGTGTGAGTGT 654
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 655 GTGCGTTTCTGGCAATGGTTCACCCCTTTCGGCTTCTGCTGCTGCTGCTGCTGCTGCT 714
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 715 GCTGATCTCTGTGGGATCATATGATCTTATC-----ATGGCTTCTCCTCAATA 765
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
DB 766 ATGCTCTGCAGCAGTGGCTCTGAGCAGACCGGAGTGTACATCATCTGTAGAGTGT--- 822
QY 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
DB 823 -----AATCTCTATAAAATTCCTAAGCTGCAGACCATG 855
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
DB 823 -----AATCTCTATAAAATTCCTAAGCTGCAGACCATG 855

Db 856 AACTATATTCGCTGGTGGGCTGCTGCTGCCAATTTTTCACACTCAGACTCTGTAT 915
Qy 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
Db 916 CTGCTGATCATTCGGTTCCTGTTAAAGTGGAGTCCCAAGATCGGGGCTGGGGTTCT 975
Qy 314 TyrIleLysAla-----ValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
Db 976 CACAGGAAGCAGCTGACCAACATCATCATCACCCTTGATCATCTCTCTTGTGTTCCTG 1035
Qy 331 ProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyrHis----AsnThrAsp 349
Db 1036 CCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAGTGGTATTATGCAAGAC 1095
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 1096 AGACTGCATAAGCTTTGGTTATACACTGCGCTTGGCAGCAGCAATGCTGCTCAAT 1155
Qy 370 ProPheLeuTyrPhe 374
Db 1156 CCTCTGCTCTATTAC 1170

RESULT 29

US-09-728-952-26
; Sequence 26, Application US/09728952
; Patent No. US20020111302A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Vicki
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020111302A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 799
; CURRENT APPLICATION NUMBER: US/09728,952
; CURRENT FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 26
; LENGTH: 1700
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (232)..(1272)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1700)
; OTHER INFORMATION: n = a,t,c or g
US-09-728-952-26

Alignment Scores:
Pred. No.: 7,29e-26 Length: 1700
Score: 378.50 Matches: 88
Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: 10 Gaps: 8

US-09-208-629F-3 (1-407) x US-09-728-952-26 (1-1700)

Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 286 CCAATGGCACCCTCAGCAATAACACAGCAGGAACTGCACAAAT-----GAA 333
Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyValPro 125
Db 334 AACTTCAAGAGAGAATTTTCCCAATGTATATCTGATAATATTTTCTGGGGAGTCTTG 393

Qy 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
Db 394 GGAATGGGTGTCATATATGTTTCTGCACCCCTTAAAGAGTCCCACTCTGTGAAC 453
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 454 GTTTTCATGTAATCTGCCATTTCCAGATCTCTCTTTCATAAGCAGCGTTCCTTCAGG 513
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 514 GGTGACTATATCTTAGAGGCTCCAAATGGATATTGGAGACCTGCGCTCGCAGGATATG 573
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 574 TCTTATTCCTTGTATGTCACATGACAGCATTTATTTCTCTGACCGTCTGAGTGT 633
Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 634 GTGCGTTTCTGGCAATGGTTCACCCCTTTCGCTTCTGCATGTCTACACGATCAGGAGT 693
Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 694 GCTGTGATCCTCTGTGGGATCATATGGATCCTTATC-----ATGGCTTCTCTCAATA 744
Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 745 ATGCTCTCGACAGTGGCTCTGAGCAGAGCGGAGTGTACATCATGCTTAGAGCTG--- 801
Qy 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu----- 281
Db 802 -----AATCTCTATAAAATGCTAAAGTGCAGACCATG 834
Qy 282 -----AlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyr 297
Db 835 AACTATATTGCTTGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
Qy 298 ThrThrLeuIleHisLysLysLysSer-----LysAspArgIleTrpLeuGly 313
Db 895 CTGCTGATCATTCGGGTCTGTTAAAGTGGAGTCCCAAGATCGGGGCTGGGGTTCT 954
Qy 314 TyrIleLysAla-----ValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
Db 955 CACAGGAAGCAGCTGACCAACATCATCATCACCCTTGATCATCTCTCTTGTGTTCCTG 1014
Qy 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHis----AsnThrAsp 349
Db 1015 CCCTATCACACTGAGGACCGTCCACTTGACGACATGGAAGTGGGTTTATGCAAGAC 1074
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 1075 AGACTGCATAAGCTTTGGTTATCACACTGGCCTTGGCAGCAGCAATGCTGCTTCAAT 1134
Qy 370 ProPheLeuTyrPhe 374
Db 1135 CCTCTGCTCTATTAC 1149

RESULT 30

US-10-225-567A-588
; Sequence 588, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Roush, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 588

; LENGTH: 2807

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-225-567A-588

Alignment Scores:

Pred. No.: 1.5e-25 Length: 2807
Score: 378.50 Matches: 88
Percent Similarity: 49.51% Conservative: 63
Best Local Similarity: 28.85% Mismatches: 121
Query Match: 17.72% Indels: 33
DB: 9 Gaps: 8

US-09-208-629F-3 (1-407) x US-10-225-567A-588 (1-2807)

```
QY 86 ProGluAspSerIleThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
DB 318 CCAATGGCACCCTTCAGCAATAACACAGCAGAACTGCACAATT-----GAA 365
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuValValGlyValPro 125
DB 366 AACTTCAAGAGAAATTTTCCCAATTGTATATCTGATAATATTTTCTGGGAGCTTG 425
QY 126 SerAsnIleValThrLeuTrpLys---LeuSerLeuArgThrLysSerIleSerLeuVal 144
DB 426 GGAATGGGTGTCCATATATGTTTCTCGCAGCTTATAAGAAGTCCACATCTGTGAAC 485
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
DB 486 GTTTTCATGCTAAATCTGGCCATTTTCAGATCTCTCTGTTCATAAGCAGCTTCCTTCAGG 545
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
DB 546 GCTGACTATATCTTAGAGCTCCAAATGATATTTGGAGACCTGGCCTGCAGGATATG 605
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
DB 606 TCTTATTCTTGTATGTCAACATGTACAGCAGATATTATTTCCTGACCGTGTGAGTGTT 665
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 666 GTCCGTTTCTCGCAATAGTTTCAACCCCTTTTCGGCTTCTGCAATGTCCACGATCAGGAGT 725
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
DB 726 GCTGTGATCTCTGTGGATCATATGATCTTATC-----ATGGCTTCTCCTCAATA 776
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
DB 777 ATGCTCTGACAGTGGCTCTGAGCAGACGCGAGTGTACATCATGTCTTAGAGCTG--- 833
QY 265 AspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeu----- 281
DB 834 -----AATCTCTATAAAATTTGCTAAGCTGCAGACCATG 866
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr 297
DB 867 AACTATATTGCTTGGTGGCTGCTCTGCTCCCAATTTTTCACATCAGCATCTGTAT 926
QY 298 ThrThrLeuIleHisLysLeuLysSer-----LysAspArgIleTrpLeuGly 313
DB 927 CTGCTGATCATTCGGGTCTGTGTTAAAGTGGAGGTCCAGAACTCGGGCTGCGGGTTCT 986
QY 314 TyrIleLysAla-----ValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
DB 987 CACAGGAAGGCACTGACCAACATCATCATCACCTTGATCATCTCTTCTGTGTTTCTG 1046
QY 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHis---AsnThrAsp 349
DB 1047 CCCTATCACACACTGAGGACCGTCCACTTGACACATGGAAAGTGGGTTTATGCAAGAC 1106
```

```
QY 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
DB 1107 AGACTGCATAAAGCTTTGGTTATCAGACTGGCTTGGCAGCAGCAATGCTGCTTCAAT 1166
QY 370 ProPheLeuTyrPhe 374
DB 1167 CCTCTGCTCTATTAC 1181
```

RESULT 31

US-10-225-567A-546
; Sequence 546, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 546
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-546

Alignment Scores:

Pred. No.: 3.85e-26 Length: 1014
Score: 378.00 Matches: 98
Percent Similarity: 50.83% Conservative: 55
Best Local Similarity: 32.56% Mismatches: 128
Query Match: 17.70% Indels: 20
DB: 9 Gaps: 9

US-09-208-629F-3 (1-407) x US-10-225-567A-546 (1-1014)

```
QY 88 AspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
DB 4 GATAAACAGGAAATCTGACAGTATCTTCTGCGACATGCCATGACACTATTGATGACTTC 63
QY 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
DB 64 CGCAATCAAGTGTATTCCACCTTGTACTCTATGATCTCTGTTGTAGGCTTCTTTGGCAAT 123
QY 128 IleValThrLeuTrpLysLeuSerLeuArgThr-----LysSerIleSerLeuValIle 145
DB 124 GGCTTTGTGCTCTATGCTCTC---ATAAAACCTATCACAGAAGTCAGCCTTCCAAAGTA 180
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 155
DB 181 TACATGATTAATTTAGCAGTAGCAGATCTACTTTGTGTGTGCACACTGCTCTCCGTGTG 240
QY 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 185
DB 241 GTCTATTATGTTTCAAAAGGCAATTTGGCTCTTTGGTGACTTCTTTGGCCGCTTCAGACCC 300
QY 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsn 205
DB 301 TATGCTTTGTATGTCACACCTCTATTGTAGCATCTCTTTTATGACACCATGAGCTTTTC 360
QY 206 ArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSer 225
DB 361 CGGTGATGCAATTTGTTTTCCAGTCCAGAACATTAATTTGTTTACACAGAAAAAGCC 420
QY 226 LeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIle 245
DB 421 AGGTTTGTGTGTAGGTATTGGATTTTGTGATTTTGCACAGTTTCTCCATTTCTAATG 480
```

QY 246 LeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAsp 265
DB 481 GCCAAACCAACAAAGAT-----GAGAAATAATACCAAGTCTTTGAGCCCCCACAA 534
QY 266 AlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGly 285
DB 535 GACAATCAAACTAAATCATATGTTTGGTCTTGCATATGTGTGTCATTG---TTTGTGGC 591
QY 286 PheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHisLysLeu--- 304
DB 592 TTTATCATCCCTTTTGTATATAATTTGCTGTACAAATGATATTTTGACCTTACTA 651
QY 305 ---LysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuVal 323
DB 652 AAAAATCAATGAAAAAATCTGTCAAGTCATAAAAGGCTATAGGAATGATGATGTC 711
QY 324 Ile-----PheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
DB 712 GTGACCGCTGCCTTTTGTAGTCAGTTCATGCCATATCATATTCACGTACCATTCAC--- 768
QY 341 AlaAsnTyrTyrTyrHisAsn-----ThrAspSerLeuTyrPheMet----- 354
DB 769 ---CTTCATTTTACACAATGAACTAAACCCCTGTGATCTGTCTTAGAATGCGAGAAG 825
QY 355 ---TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyr 373
DB 826 TCGTGTGTACATACCTTGTCTGTGGTGCATCAATTTGCTTTGACCTCTCTCTATAT 885
QY 374 Phe 374
DB 886 TTC 888

RESULT 32

US-10-225-567A-218
; Sequence 218, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glemma C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 218
; LENGTH: 2424
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-218

Alignment Scores:
Pred. No.: 1.87e-25 Length: 2424
Score: 376.50 Matches: 95
Percent Similarity: 47.44% Conservative: 72
Best Local Similarity: 26.99% Mismatches: 130
Query Match: 17.63% Indels: 55
DB: 9 Gaps: 12

US-09-208-629F-3 (1-407) x US-10-225-567A-218 (1-2424)

QY 46 AlalysProThrLeuThrIleLysSerPheAsnGlyProGlnAsnThrPheGluGlu 65
DB 159 GCTGTCCCAACGCGGACGGAGCGTCTCTGTGCGCGTCCGGGTTCGTCC----- 209
QY 66 PheProLeuSerAspIleGluGlyTyrThrGlyAlaThrThrThrIleLysAlaGluCys 85
DB 210 -----TGGGGGAACAGCAGCGTCGCC----- 230

QY 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAla---ThrIleGlyTyrLeuArg 104
DB 231 -----TCCACTGCCCGCTCTCTCGTCTCAAAATGCGCCTTGACCAAG 275
QY 105 SerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyVal 124
DB 276 ACGGCGTTCACAGTCTTACTACTGCGCGTGTCTACATCTTGGTATTTCATCATCGGCTTC 335
QY 125 ProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLys-----SerIle 141
DB 336 CTGGGCAACAGCGTGGCCATCTGGATGTCTGCTTCCCATGAAAGCCCTGGAGCGGCATC 395
QY 142 SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeu 161
DB 396 TCCGTGTACATGTC-----AATTGGCTCTGGCCGACTTCTTGTACGTGTGACTCTG 449
QY 162 PropheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCys 181
DB 450 CCAGCGCTGATCTTCTACTTCAATAAAACAGACTGGATCTTCGGGATGCCATGTGT 509
QY 182 ArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCys 201
DB 510 AAATGCGAGAGGTTCATCTTTCATGTGAACCTCTATGGCAGCACTTGTCTTGACATGC 569
QY 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys 221
DB 570 ATCAGTGGCCACCGGTACAGCGGTGTGTGTATCCCTCAAGTCCCTGGCGCGCTCAAA 629
QY 222 ArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeu 241
DB 630 AAGAAAGATGCGATCTGTATCAGCGTGTGTGTGCTCATTTGTGTGGTGTGGTATCTCC 689
QY 242 PropheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHis 261
DB 690 CCC---ATCCTCTTCTACTCAGTACCGGGGTCGCAAAACAAACACCATCCTGTTAC 746
QY 262 AspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeu 281
DB 747 GACCCACC-----TCAGACGAGTACCTCGGAAGTTATTTTCATCTACAGCATG 794
QY 282 -----AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThr 299
DB 795 TGCACGACCGTGGCCATGTTCTGTGTCCTTGGTGTGATCTTGGGCTGTTACGGATTA 854
QY 300 LeuIleHisLysLeuLysSerLysAsp-----ArgIle 310
DB 855 ATTTGAGAGCTTTGATTTACAAAGATCTGGACAACCTCTCTCTGAGGAGAAATCGATT 914
QY 311 TrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
DB 915 TACCTG-----GTAATCATTTGACTGACTGTTTTCGTGCTGTACATC 959
QY 331 ProThrAsnIleLeuValIleHis---HisAlaAsnTyrTyrTyrHisAsn----- 347
DB 960 CTTTTCATGTGATAAAACGATGAACCTTGAGGGCGCGCTTGAATTTTCAGACCCCGACA 1019
QY 348 -----ThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer 363
DB 1020 ATGTGTGCTTCAATGACAGGTTTATGCCAGTATCAGGTGACAAAGAGGCTTAGCAAGT 1079
QY 364 LeuAsnSerCysLeuAspProPheLeuTyrPheVal 375
DB 1080 CTCAACAGTTGTGGAGCCCATCTCTATTCTTG 1115

RESULT 33

US-10-167-192-4
; Sequence 4, Application US/10167192
; Publication No. US20030040052A1
; GENERAL INFORMATION:
; APPLICANT: Gluckmann, Maria A.
; TITLE OF INVENTION: No. US20030040052A1el G-Protein Coupled Receptors
; FILE REFERENCE: 5800-11A
; CURRENT APPLICATION NUMBER: US/10/167,192


```
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: US/09/420,187
; PRIOR FILING DATE: 1999-10-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1358
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-167-192-4

Alignment Scores:
Pred. No.: 1,39e-25 Length: 1358
Score: 374.00 Matches: 112
Percent Similarity: 47.14% Conservative: 61
Best Local Similarity: 30.52% Mismatches: 144
Query Match: 17.51% Indels: 50
DB: 9 Gaps: 14

US-09-208-629F-3 (1-407) x US-10-167-192-4 (1-1358)

QY 42 SerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
DB TCTGAGAACAAACGCAAGACAGTA-----AACTGGACCTTGAAGT 109
QY 62 ThrPheGluGluPheProLeu-----SerAspIleGluGly 73
DB 110 TAGCAGCGTGGGCTTCTCTAATATTACACCGTAAAGGCAATTCATCACCATAAGAAGGA 169
QY 74 TrpThrGlyAlaThrThrIleLysAla---GluCysProGlu-----87
DB 170 -----ACATTGTGAAGGTACTCCAGTGCAGAAAGAGGACAAAGCAGA 214
QY 88 -----AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGly 101
DB 215 CATTCGTAGAGAACATGGATGAACAGGAAATCTGCAGTATCTCTGCCACATGCCAT 274
QY 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheVal 121
DB 275 GACACTATTGATGACTTCGCAATCAAGTGTATTCACCTGTGTACTGTATGATCTCTGTT 334
QY 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr-----Lys 139
DB 335 GTAGGCTCTTTGGCAATGGCTTTGTGCTCTATGTGCTC---ATAAAACCTATCAAG 391
QY 140 SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysVal 159
DB 392 AAGTCAGCCTTCCAAGTATACATGATTAATTTAGCAGTAGCAGATCTACTTTGTGTGTC 451
QY 160 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 179
DB 452 ACACTGCCTCTCCGTGTGGTCTATTATGTTTCAAAAGGCATTTGGCTCTTTGGTACTTC 511
QY 180 MetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeu 199
DB 512 TTGTGCGCGCTCCAGCCTATGCTTTGTGTGTAGGTATTTGGATTTTGTGATTTTGACC 571
QY 200 ThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeu 219
DB 572 ACAGCCATGAGCTTTTCCGCTGCAATGCAATTTGTTTCCAGTCCAGAACATAATTG 631
QY 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr 239
DB 632 GTTACACAGAAAAAGCCAGGTTTGTGTGTAGGTATTTGGATTTTGTGATTTTGACC 691
QY 240 MetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThr 259
DB 692 AGTTCTCCATTTCTAATGTGGCCAAACACAAAAAGAT-----GAGAAAAATAATCAAC 745
QY 260 CysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheVal 279
DB 746 TGCTTTGAGCCCCCACAAGACAATCAAACTAAATAATCATGTTTGTGCTTGGCATTTATGTG 805

; Sequence 1, Application US/10023775B
; Publication No. US20030022282A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Ltd. (EP, GB, JP, except GB)
; APPLICANT: Pfizer Inc. (US, JP, except GB)
; APPLICANT: Fidoock, Mark David
; TITLE OF INVENTION: No. US20030022282A1 Polypeptide
; FILE REFERENCE: PC10959AGPR
; CURRENT APPLICATION NUMBER: US/10/023,775B
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: GB 0030854.4
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,590
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/296,660
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0111031.1
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-775B-1

Alignment Scores:
Pred. No.: 2.4e-25 Length: 1014
Score: 369.50 Matches: 87
Percent Similarity: 49.32% Conservative: 58
Best Local Similarity: 29.59% Mismatches: 108
Query Match: 17.30% Indels: 41
DB: 9 Gaps: 8

US-09-208-629F-3 (1-407) x US-10-023-775B-1 (1-1014)

QY 107 LeuSerThrGlnValIleProAlaIleTyrIleLeuPheValValGlyValProSer 126
DB 91 CTCAGATGCATCTACCTCCCTGTTATTATTTATGTCATTTATCTCTCTGGGATTTCCAGGC 150
QY 127 AsnIleValThrLeu-----TrpLysLeuSerLeuArg 137
DB 151 AATGCAGTAGTAGTATCCATTTTACATTTTCAAAATGAGACCTTGAAG-----198
QY 138 ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe 157
DB 199 -----AGCAGCACCATCATTTATGCTGTAACCTGGCCTGCAGATCTGCTGTAT 246
```



```
; APPLICANT: Tian, Hui
; APPLICANT: Dai, Kang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Zhao, Jiagang
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030082585A1el Receptors
; FILE REFERENCE: 018781-008410US
; CURRENT APPLICATION NUMBER: US/10/188, 405
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 60/302,800
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human TGR164
US-10-188-405-7

Alignment Scores:
Pred. No.: 2,4e-25 Length: 1014
Score: 369.50 Matches: 87
Percent Similarity: 49.32% Conservative: 58
Best Local Similarity: 29.59% Mismatches: 108
Query Match: 17.30% Indels: 41
DB: Gaps: 8

US-09-208-629F-3 (1-407) x US-10-188-405-7 (1-1014)
QY 107 LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer 126
Db 91 CTCAGATGACGATACCTCCCTGTTATTTATGTCATTTCTCTCGGGATTCAGGC 150
QY 127 AsnIleValThrLeu-----TrpLysLeuSerLeuArg 137
Db 151 AATGCGATGATGATACCTTCAAAATGAGACCTTGGGAA----- 198
QY 138 ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe 157
Db 199 -----AGCAGCACCATTCATTGCTGACCTGGCCTGCACATCTGCTGTAT 246
QY 158 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 177
Db 247 CTCAGCAGCTCCCTTCCTGATTCATCTACTATGCCAGTGGCGAAACTGGATCTTGA 306
QY 178 GluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu 197
Db 307 GATTTCATGTGAAGTTTATCCGCTTCAGCTTCCATTTCAACCTGTATAGCAGCATCCTC 366
QY 198 IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaIleAlaHisProPheThrTyrGln 217
Db 367 TTCCTCACCTGTTTCAGCATCTCCGCTACTGTGTGATCATTCACCCATGACCTGCTTT 426
QY 218 LysLeuProLysArgSerPheSerLeuMetCysGlyIleValTrpValMetValPhe 237
Db 427 TCATTCACAAACTCGATGTCAGTGTAGCTGTGCTGTGTGTGGATCATTTCACTG 486
QY 238 LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle 257
Db 487 GTAGTGTCATTCGGATGACCTTCTTGATCACATCAACCAAC-----AGCACCAACAGA 540
QY 258 ThrThrCysHisAspValAlaCysGluSerProSerSerPheArgPheTyr 277
Db 541 TCAGCTGTCTCAGCTCACCAGTTCGGATGAATC---AATACTATTAGTGGTGAAC 597
QY 278 PheValSerLeuAlaPhePheGlyPheLeuIlePheValIleIlePheCysTyr 297
Db 598 CTGATT---TTGACTGCAACTACTTTCGCTCCCTTGTGTGTATGACACTTTGCTAT 654
QY 298 ThrThrLeuIleHisLys-----LeuLysSerLys 307
```

```
Db 655 ACCACGATTATCCACACTCTGACCCCATGACTGCAAACTGACAGCTGCCTTAAGCAGAA 714
QY 308 AspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIle 327
Db 715 GCACGA-----AGGTAACCATCTGCTACTCTTCGATTTTACGTA 756
QY 328 CysPheAlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsn 347
Db 757 TGTITTTTACCTTCATATCTTGGAGGTTCATTCGGATCGAATCTCGCCTGCTTCAATC 816
QY 348 Thr-----AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer 363
Db 817 AGTTGTTCATTGAGAATCAGATCCATGAAGCTTACATCGTTTCTAGACCATTAGCTGCT 876
QY 364 LeuAsnSerCysLeuAspPhePheLeuTyrPheValMetSer 377
Db 877 CTGACACACCTTTGGTAACCTGTACTATATGTGGTGCAGC 918

RESULT 37
US-09-885-453-2
; Sequence 2, Application US/09885453
; Publication No. US2003008080A1
; GENERAL INFORMATION:
; APPLICANT: Communi, Didier
; TITLE OF INVENTION: RECEPTOR GPCRx10
; FILE REFERENCE: 9409/2082
; CURRENT APPLICATION NUMBER: US/09/885,453
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/885,453
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DNA nucleotide sequence
; LOCATION: (1)..(1014)
; OTHER INFORMATION: GPCRx10 DNA sequence
US-09-885-453-2

Alignment Scores:
Pred. No.: 2,4e-25 Length: 1014
Score: 369.50 Matches: 87
Percent Similarity: 49.32% Conservative: 58
Best Local Similarity: 29.59% Mismatches: 108
Query Match: 17.30% Indels: 41
DB: Gaps: 8

US-09-208-629F-3 (1-407) x US-09-885-453-2 (1-1014)
QY 107 LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer 126
Db 91 CTCAGATGACGATACCTCCCTGTTATTTATGTCATTTCTCTCGGGATTCAGGC 150
QY 127 AsnIleValThrLeu-----TrpLysLeuSerLeuArg 137
Db 151 AATGCGATGATGATACCTTCAAAATGAGACCTTGGGAA----- 198
QY 138 ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe 157
Db 199 -----AGCAGCACCATTCATTGCTGACCTGGCCTGCACATCTGCTGTAT 246
QY 158 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 177
Db 247 CTCAGCAGCTCCCTTCCTGATTCATCTACTATGCCAGTGGCGAAACTGGATCTTGA 306
QY 178 GluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu 197
Db 307 GATTTCATGTGAAGTTTATCCGCTTCAGCTTCCATTTCAACCTGTATAGCAGCATCCTC 366
QY 198 IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaIleAlaHisProPheThrTyrGln 217
Db 367 TTCCTCACCTGTTTCAGCATCTCCGCTACTGTGTGATCATTCACCCATGACCTGCTTT 426
QY 218 LysLeuProLysArgSerPheSerLeuMetCysGlyIleValTrpValMetValPhe 237
Db 427 TCATTCACAAACTCGATGTCAGTGTAGCTGTGCTGTGTGTGGATCATTTCACTG 486
QY 238 LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle 257
Db 487 GTAGTGTCATTCGGATGACCTTCTTGATCACATCAACCAAC-----AGCACCAACAGA 540
QY 258 ThrThrCysHisAspValAlaCysGluSerProSerSerPheArgPheTyr 277
Db 541 TCAGCTGTCTCAGCTCACCAGTTCGGATGAATC---AATACTATTAGTGGTGAAC 597
QY 278 PheValSerLeuAlaPhePheGlyPheLeuIlePheValIleIlePheCysTyr 297
Db 598 CTGATT---TTGACTGCAACTACTTTCGCTCCCTTGTGTGTATGACACTTTGCTAT 654
QY 298 ThrThrLeuIleHisLys-----LeuLysSerLys 307
```

Qy	198	IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaHisProPheThrTyrGln	217
Dd	367	TTCCCTCACCTGTTCAGCATCTTCGGCTACTGTGTGATCATTCACCAATGACGTCTTT	426
Qy	218	LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPhe	237
Dd	427	TCCATTACAAAACTCGATGTGCAGTTGTACGCTTGCTGTGGTGTGGATCAITTCACGTG	486
Qy	238	LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluile	257
Dd	487	GTAGCTGTTCATTCGGATGACCTTCTTGATCATCAACCAAAC-----AGGACCAACA	540
Qy	258	ThrThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyr	277
Dd	541	TCAGCCTGTCTCGACCTCACCAGTTTCGATGAACTC---AATACTATTAAAGTGGTACAAC	597
Qy	278	PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr	297
Dd	598	CTGATT---TTGACTGCCAACTACTTTCGCTCCCTCCCTGGTGATAGACACTTTCGTAT	654
Qy	298	ThrThrLeuIleHisLys-----LeuLysSerLys	307
Dd	655	ACCACGATTATCCACACTCTGACCCATGGAGCTGCAAACTGACAGCTGCCCTTAAGCAGAAA	714
Qy	308	AspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIle	327
Dd	715	GCACGA-----AGGCTAACCACTTCTGTACTCTCTTGCATTTTACGTA	756
Qy	328	CysPheAlaproThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHisAsn	347
Dd	757	TGTTTTTACCCCTTCATATCTTAGGGGTCAATCGGATCGAATCTCGCCGTGTTTCAATC	816
Qy	348	Thr-----AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer	363
Dd	817	AGTGTGTCCATTGAGNAATCAGATCATGAAGCTTACATCGTTCTAGACCAATTAGCTGCT	876
Qy	364	LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSer	377
Dd	877	CTGAACACCTTTGGTAACCTGTTATCTATATGTGGTGGTGCAGC	918

RESULT 38

RESULT 38
US-10-079-384-13
; Sequence 13, Application US/10079384
; Publication No. US20030108986A1
; GENERAL INFORMATION:
; APPLICANT: Communi, Digier
; TITLE OF INVENTION: COMPOSITIONS AND METHODS COMPRISING G-PROTEIN COUPLED RECEPTORS

Alignment Scores:		
Pred. No.:	2.4e-25	1014
Score:	369.50	87
Percent Similarity:	49.32%	Conservative: 58
Best Local Similarity:	29.59%	Mismatches: 108
Query Match:	17.30%	Indels: 41
DB:	9	Gaps: 8

US-09-208-629F-3 (1-407) x US-10-079-384-13 (1-1014)

Qy	107	LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer	129
Db	91	CTCAAGATGCACCTACCTCCCTGTTATTATGGCATTAATCTTCCTCGTGGGATTCACAGC	150
Qy	127	AsnIleValThrLeu-----TTPLSLeuSerLeuArg	137
Db	151	AATCGAGTAGTGATATCCACTTACATTTTCAAAATGAGACCTTGGAG-	198
Qy	138	ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe	157
Db	199	-----AGCAGCACCATCATTTATGCTGAACCTGGCGCTGCACAGATCTGCTGTAT	246
Qy	158	CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly	177
Db	247	CTGACCAGCCCTCCCTGATTACACCTACATGCGAGTGGCGAAATCTGGATCTTTGGA	306
Qy	178	GluValMetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu	197
Db	307	GATTTCATGTTAGTTTATCCGCTTCAGCTTCATTTCAACCTGTATAGCAGCATCTCTC	366
Qy	198	IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln	217
Db	367	TTCTCACCTGTTTTCAGCATCTTCGCTACTGTGTGATCATTTACCCCAATGAGCTGCTT	426
Qy	218	LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPhe	237
Db	427	TCCATTACAAAACTCGATGTCGAGTTGTAGCCTGTGCTGGTGTGATCATTTCAAC	486
Qy	238	LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle	257
Db	487	GTAGCTGTTCATTCGATGACCTTCTTGATCATCATCAACCAAC-----AGGACCAACAGA	540
Qy	258	ThrThrCysHisAspValValAlaCysGluSerProSerSerPheArgPheTyrTyr	277
Db	541	TCAGCTGTCTCGACCTCACCAGTTTCGGATGAATC---AATACTATTAAAGTGTACAA	597
Qy	278	PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr	297
Db	598	CTGATT---TTGATCGCAACTACTTTCTCGCCTCCCTTGGTGATAGTACACTTTGCTAT	654
Qy	298	ThrThrLeuIleHisLys-----LeuLysSerLys	307
Db	655	ACCAGATTATCCACCTCTGACCCATGGATCGAAACTGACAGCTGCCTTAAGCAGAAA	714
Qy	308	AspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIle	327
Db	715	GCAGCA-----AGGCTAACCAATCTGCTACTCTCTGCTGCTTACGTA	756
Qy	328	CysPheAlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsn	347
Db	757	TGTTTTTACCCTTCCCATCTTGGAGTGCATTCGGATCGAACTCGCCTGCTTCAATC	816
Qy	348	Thr-----AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer	363
Db	817	AGTTCTTCCATTGAGAATCAGATCCATGAGCTTACATCGTTTCTTAGACCATAGCTGCT	876
Qy	364	LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSer	377
Db	877	CTGAACACCTTGGTAACTGTTACTATATGTGTGGTGCTCAGC	918

RESULT 39
 US-10-225-567A-646
 ; Sequence 646, Application US/10225567A
 ; Publication No. US20030113798A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LifeSpan Biosciences
 ; APPLICANT: Brown, Joseph P.
 ; APPLICANT: Burmer, Glenn A.
 ; APPLICANT: Roush, Christina L.
 ; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED
 ; FILE REFERENCE: 1920-4-4
 ; CURRENT APPLICATION NUMBER: US/10/225-567A

ORGANISM: Homo sapiens
US-10-225-567A-646

; GENERAL INFORMATION:
 ; APPLICANT: Glaxo Group Limited
 ; TITLE OF INVENTION: Polypeptide
 ; FILE REFERENCE: OG1021
 ; CURRENT APPLICATION NUMBER: US/09/943,798
 ; CURRENT FILING DATE: 2001-08-31

RESULT 40
US-09-943-798-3
; Sequence 3, Application US/09943798
; Patent No. US20020065215A1
; GENERAL INFORMATION:
; APPLICANT: Glaxo Group Limited
; TITLE OF INVENTION: Polypeptide
; FILE REFERENCE: QG1021
; CURRENT APPLICATION NUMBER: US/09/943,798
; CURRENT FILING DATE: 2001-08-31

US-09-208-629F-3 (1-407) X US-09-943-798-3 (1-1014)

Qy	107	LeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSer	126
Db	91	CTCAAGATGCATACCTCCCTGTTATTATGGCATATCTCTCGTGGGATTTCCAGGC	150
Qy	127	AsnIleValThrLeu-----TrpIysLeuSerLeuArg	137
Db	151	AATGCAGTAGTGATATCCACTTACATTTTCAAAATGAGACTTGGAG-----	198
Qy	138	ThrLysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPhe	157
Db	199	-----AGCAGCACCATCATTTATGCTGAACCTGGCGTCACAGATCTGCTGAT	246
Qy	158	CysValThrLeuProPheIysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly	177
Db	247	CTGACCAAGCTCCCTTCCTGATTCACCTACTACTGCCAGTGGCGAAAACTGGATCTTTGGA	306
Qy	178	GluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeu	197
Db	307	GATTTCATGTGAAGTTTATCGGCTTCAGCTTCCATTTCAACCTGTATAGCAGATCCCTC	366
Qy	198	IleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGln	217
Db	367	TTTCCTCACCTGTTTCAGCATCTTCGGTACTGTGTGATCATTCACCAATGAGCTGCTTT	426
Qy	218	LysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPhe	237
Db	427	TCCATTACAAAACTCGATGTGCAGTTGTAGCCCTGTGCTGTGGTGTGGATCATTTCACTG	486
Qy	238	LeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIle	257
Db	487	GTAGCTGTCTATCCGATGACCTTCTTGATCATCATCAACCAAC-----AGGACCAACAGA	540
Qy	258	ThrThrCysHisAspValValAspAlaCysGluSerProSerPheArgPheTyrTyr	277
Db	541	TCAGCCTGTCTCGACCTCCACCGATTTCGATGAATC---AATACTATTAGTGGTCAAC	597
Qy	278	PheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyr	297

```
Db 598 CTGATT---TTGACTGCAACTACTTTCTGCTCCCTTGTGTAGTGACACTTTGCTAT 654
Qy 298 ThrThrLeuIleHisLys-----LeuLysSerLys 307
Db 655 ACCACGATTATCCACACTCTGACCCATGGAGTGCAAACTGACAGCTGCCTTAAGCAGAAA 714
Qy 308 AspArgIleTrpLeuGlyTyriLeLysAlaValLeuLeuIleLeuValIlePheThrIle 327
Db 715 GCAGCA-----AGGCTAACCATTCGTACTCTCTTGCATTTACGTA 756
Qy 328 CysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyriTyriHisAsn 347
Db 757 TGTITTTTACCCCTTCCATATCTTGAGGGTCATTGGAGTCGAATCGCTGCTTTCAATC 816
Qy 348 Thr-----AspSerLeuTyriPheMetTyriLeuIleAlaLeuCysLeuGlySer 363
Db 817 AGTTGTTCCATTGAGAAATCAGATCCATGAGCTTACATCGTTTCTAGACCATTAGCTGCT 876
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyriPheValMetSer 377
Db 877 CTGACACACCTTGGTAACTGTTACTATATGTGTGTGTGCTCAGC 918
```

Search completed: June 29, 2003, 11:46:16
Job time : 170.838 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:52:59 ; Search time 45.9436 Seconds
(without alignments)
2716.756 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLTXQHPVAGSQDKMKIL.....AMARPLXRRPRDIWEDIHAW 407

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09208629/runat 24062003 101626 12944/app query.fasta_1.1166
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629 @CIGN 1 1 47 @runat 24062003 101626 12944 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2121	99.3	1224	2	US-08-742-440A-1
2	1443.5	67.6	1224	2	US-08-742-440A-4
3	1369	64.1	1102	2	US-08-742-440A-5
4	1300.5	60.9	1124	2	US-08-742-440A-2
5	584	27.3	2732	1	US-08-476-000-60
6	584	27.3	2732	1	US-08-472-840-60
7	584	27.3	2732	2	US-08-476-976-60
8	584	27.3	2732	4	US-08-474-410-60
9	584	27.3	2732	4	US-08-486-673B-60
10	579.5	27.1	1475	1	US-08-097-938-1
11	579.5	27.1	1475	1	US-08-476-000-1
12	579.5	27.1	1475	1	US-08-472-840-1

13	579.5	27.1	1475	2	US-08-476-976-1
14	579.5	27.1	1475	3	US-08-474-410-1
15	579.5	27.1	1475	4	US-08-486-673B-1
16	563	26.4	1255	1	US-08-097-938-3
17	563	26.4	1255	1	US-08-476-000-3
18	563	26.4	1255	1	US-08-472-840-3
19	563	26.4	1255	2	US-08-476-976-3
20	563	26.4	1255	3	US-08-474-410-3
21	563	26.4	1255	4	US-08-486-673B-3
22	551	25.8	1414	1	US-08-476-000-62
23	551	25.8	1414	1	US-08-472-840-62
24	551	25.8	1414	2	US-08-476-976-62
25	551	25.8	1414	3	US-08-474-410-62
26	551	25.8	1414	4	US-08-486-673B-62
27	511.5	23.9	1764	1	US-08-313-553-12
28	511.5	23.9	1764	3	US-08-767-993-12
29	509	23.8	3480	1	US-07-657-769B-68
30	509	23.8	3480	1	US-07-789-184-219
31	509	23.8	3480	1	US-08-475-263-219
32	509	23.8	3480	1	US-08-485-886-219
33	509	23.8	3480	2	US-08-477-362-219
34	509	23.8	3480	2	US-08-477-134-219
35	509	23.8	3480	3	US-08-473-489A-219
36	509	23.8	3480	3	US-08-485-695-219
37	509	23.8	3480	4	US-08-018-760-219
38	507.5	23.8	4895	3	US-09-053-866-1
39	507.5	23.8	4895	4	US-09-479-130-1
40	431.5	20.2	1643	1	US-08-383-750-3
41	431.5	20.2	1643	3	US-08-352-678-3
42	431.5	20.2	1643	5	PCT-US93-09636-3
43	398.5	18.7	1901	1	US-08-153-848-43
44	398.5	18.7	1901	3	US-09-299-843A-43
45	398.5	18.7	1901	4	US-09-088-337B-43

ALIGNMENTS

RESULT 1
US-08-742-440A-1
; Sequence 1, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihara, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400

TELEFAX: 650 327-3231

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1224 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-742-440A-1

Alignment Scores:

Pred. No.: 2,43e-211 Length: 1224

Score: 2121.00 Matches: 404

Percent Similarity: 99.26% Conservative: 0

Best Local Similarity: 99.26% Mismatches: 3

Query Match: 99.30% Indels: 0

DB: 2 Gaps: 0

US-09-208-629F-3 (1-407) x US-08-742-440A-1 (1-1224)

Qy 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20

Db 3 ACTTTGTATACTTAACAAACATCTCTAGCCGGGTCTCAGGACATCAAGATGAAATCCTT 62

Qy 21 IleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsn 40

Db 63 ATCTTGGTTCAGCTGGGGTCTGCTTTCTGCCAGTCACTGTTCGCCAAAGTGCATAAAT 122

Qy 41 ValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGln 60

Db 123 GTTTTCAGCAACTCAGCAAGCCAACTTAACATTAAAGAGTTTAAATGGGGTGCCTCAA 182

Qy 61 AsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThr 80

Db 183 AATACCTTTGAAGAAATTCCTCACTTCTGACATAGAGGGCTGGACAGAGCCACCAACT 242

Qy 81 IleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAlaThrIle 100

Db 243 ATAAAGCGAGTGTCCGAGGACAGATTTCACTCTCCAGTGAATAATGTACCATTA 302

Qy 101 GlyTyrLeuArgSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPhe 120

Db 303 GGATACCTGAGAAGTTCCTTAAGTACCAAGTATACCTGCCATCTATATCTGCTGTT 362

Qy 121 ValValGlyValProSerAsnIleValThrLeuThrLysLeuSerLeuArgThrLysSer 140

Db 363 GTGGTTGGTGTPACCATCCAACTCGTACCTGTGGAACTCTCTTAAAGGACCAATCC 422

Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160

Db 423 ATCAGTCTGGTCATCTTTCACCACTGGCCATCGCAGATCTCTTTTCTGTGCACA 482

Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180

Db 483 CTGCCATTTAAGATCGCCTACCATCTCAATGGCAACAACTGGGTATTTGGCAGGTCA 542

Qy 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200

Db 543 TGCCGGATCACCACGGTCTGTTTCTAGCGCAACATGACTGCGCTATCTCTGATCCTCA 602

Qy 201 CysMetGlyIleAsnAtgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220

Db 603 TGCATGGGCATCAACCGCTACTGGCCAGGGCTCACCTTTCACATACCAAGAGTCC 662

Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240

Db 663 AAACGCAGCTTCTCTTGTCTCATGTGTGCATAGTGTGGGTCACTGTTTCTTATACAT 722

Qy 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCys 260

Db 723 CTGCCCTTTGTTCATCTCTGAAGCAGGAGTACCACCTCGTGCCTCAGAGATCACCCAT 782

Qy 261 HisAspValValAspAlaCysGluSerProSerPheArgPheTyrThrPheValSer 280

Db 783 CACGATGTCGTCGACGGTGGAGTCCCATCATCTCTCCGATTTCTACTACTTCGTCCTCC 842

Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300

Db 843 TTAGCATTTCTTTGGGTTCCTCATCCGTTTGTGATCATCATCTTCTGTGTACAGACTCTC 902

Qy 301 IleHisLysLeuLysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeu 320

Db 903 ATCCAAAACCTTAAATCAAGGATCGGATATGGTGGGTACATCAAGCCGCTCTCTCTC 962

Qy 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340

Db 963 ATCTTGTGATTTTCAATTTTGCCTTTCCTCCCAACATCATCTCTGTAATCCACCAT 1022

Qy 341 AlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleLeuAlaLeuCys 360

Db 1023 GCCAATCTACTACCAATACCCAGACAGCTTGTATCTTATGTATCTTATGCTCTGTGC 1082

Qy 361 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValVal 380

Db 1083 CTGGGAGCTGAATAGCTGCTAGATCCATCTCTTACTTCTGTCATGTCGAAAGTTGTA 1142

Qy 381 AspGlnLeuAsnPro***SerAlaMetAlaArgProLeu***ArgProArgArgAspIle 400

Db 1143 GATCAGCTTAATCTCTTAGTCGCAATGGCAAGACCACTTTAGAGACCAAGGAGATATC 1202

Qy 401 TrpGluAspIleHisAlaTrp 407

Db 1203 TGGGAAGACATACATGCTTGG 1223

RESULT 2

US-08-742-440A-4

; Sequence 4, Application US/08742440A

; Patent No. 5892014

; GENERAL INFORMATION:

; APPLICANT: Coughlin, Shaun

; APPLICANT: Ishihari, Hiroaki

; APPLICANT: Comolly, Andrew

; TITLE OF INVENTION: Protease Activated Receptor

; TITLE OF INVENTION: 3 and Uses Thereof

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bozicevic & Reed, LLP

; STREET: 285 Hamilton Avenue, Suite 200

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/742,440A

; FILING DATE: 30-OCT-1996

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Sherwood, Pamela J

; REGISTRATION NUMBER: 36,677

; REFERENCE/DOCKET NUMBER: UCAL/060PAT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-327-3400

; TELEFAX: 650 327-3231

; TELEX:

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1224 base pairs


```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-742-440A-4

Alignment Scores:
Pred. No.: 6,84e-141 Length: 1224
Score: 1443.50 Matches: 264
Percent Similarity: 83.74% Conservative: 45
Best Local Similarity: 71.54% Mismatches: 59
Query Match: 67.58% Indels: 1
DB: 2 Gaps: 1

US-09-208-629F-3 (1-407) x US-08-742-440A-4 (1-1224)

Qy 11 GlySerGlnAspIleLeuValLeuValAlaAlaGlyLeuLeuPheLeu 30
Db 40 GGGACTCAGGTCAAAATGAAGCCCTCATCTTTCAGCTGCTGGCCCTCTCTG 99
Qy 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspSerAlaLysProThr 49
Db 100 TTGCCCACTTTTGTGTCAGATGGATGGAATAATGATACAAACACTTGGCCAAAGCCAAAC 159
Qy 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
Db 160 TTACCCATTAAAGACTTTCGTGGAGCTCCCCCAAATCTTTTGAAGAGTTCCCTTTTCT 219
Qy 70 AspileGluGlyTrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSer 89
Db 220 GCCTTGAAGGCTGGAGCAGGACCATGATTACTGTAAATAATTAAGTCCCTGGAAGAAGT 279
Qy 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThr 109
Db 280 GCCTTCACATCCCATGTGAATAATGCTACATGGGGTACTGACCACCTCTTAAAGTACT 339
Qy 110 GlnValIleProAlaIleTyrIleLeuPheValValGlyValProSerAsnIleVal 129
Db 340 AAATGATACCTGCATCTACCTCTGCTGTTGTAGTGTGTCGGCCCAATGCTGTG 399
Qy 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149
Db 400 ACCCTGGAGTCTTTCTTCAGAGCAGATCCATCTGTACCACTGTATTCTACACCAAC 459
Qy 150 LeuAlaIleAlaAspLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 169
Db 460 CTGCCCATTCAGATTTCTTTTGTGTACATGGCCCTTAAAGATAGTTATCATCTC 519
Qy 170 AsnGlyAsnThrValPheGlyGluValMetCysArgIleThrThrValValPheTyr 189
Db 520 AATGGGAACAATGGGTATTGGAGAGGTCTGTGCGGGCCACACAGTCACTCTCTAT 579
Qy 190 GlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAla 209
Db 580 GGCAACATGATCTCTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 639
Qy 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229
Db 640 ATCGTCCATCTTTACCTACCGGGGCTGCCAAGCAGACACCTATGCTTGGTAACTAT 699
Qy 230 GlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlu 249
Db 700 GGACTGGTGGGCAACAGATTTTCTATATATGCTGCCATTTTTCATCTAGTGAAGCAGAA 759
Qy 250 TyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSer 269
Db 760 TATTATCTTGTTCAGCCAGACATCACCACCTGCCATGATGTTCACAACTTGGAGTCC 819
Qy 270 ProSerSerPheArgPheTyrTyrPheValSerLeuAlaPheGlyPheLeuIlePro 289
Db 820 TCATCTCCCTTCCAACTCTATTACTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 879
Qy 290 PheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerLysAspArg 309
```

```
880 TTTGTCTTATCATCTACTGTCATGCGGACCATCATCGGACACTTAATGATCATGATCAT 939
310 IleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIleCysPhe 329
940 AGATGGTGTGGTATGTTAAGGGAGTCTCTCATCTCTGTGATTTTACCATTTGCTTT 999
330 AlaProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAsp 349
1000 GCTCCAAGCAATATTATTCTTATTATTCACCATGCTAACTACTACTACAACAACACTGAT 1059
350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
1060 GGCTTATATTTATATATCTCATAGCTTTGTGCTGGTAGTCTTAATAGTTGCTTAGAT 1119
370 ProPheLeuTyrPheValMetSerLys 378
1120 CCATTCCTTATTCTTCATGCAAAA 1146

RESULT 3
US-08-742-440A-5
; Sequence 5, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1102 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-742-440A-5

Alignment Scores:
Pred. No.: 3.28e-133 Length: 1102
Score: 1369.00 Matches: 245
Percent Similarity: 84.96% Conservative: 43
Best Local Similarity: 72.27% Mismatches: 51
Query Match: 64.09% Indels: 0
DB: 2 Gaps: 0
```

```
US-09-208-629F-3 (1-407) x US-08-742-440A-5 (1-1102)
Qy 40 AsnValSerAspSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyPro 59
Db 13 AATGATACAACTTGGGAAGCCAACTTACCCATAAGACCTTTCGTGAGCTCC 72
Qy 60 GlnAsnThrPheGluGlyPheProLeuSerAspIleGluGlyTrpThrGlyAlaThr 79
Db 73 CCAATTTCTTTGAAGAGTTCCTTTCTGCTTGAAGGCTGGACAGGACGATT 132
Qy 80 ThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAlaThr 99
Db 133 ACTGTAATAATTAAGTCCCTGGAAGAGTCTTACATCTCCATGTGAAAAATGCTACC 192
Qy 100 IleGlyTrpLeuArgSerSerLeuSerThrGlnValIleProAlaIleThrIleLeu 119
Db 193 ATGGGGTACCTGACCACTCTTAAGTACTAAATGATGATGATGATGATGATGATG 252
Qy 120 PheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLys 139
Db 253 TTGTAGTGTGTCCTGGGCAATGCTGTGACCTGTGGATGCTTTTCTTCAGGACGAGA 312
Qy 140 SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysVal 159
Db 313 TCCATCTGTACCACTGATTCTACCACTGSCCATTCGAGATTTCCTTTTGTGT 372
Qy 160 ThrLeuProPheLysIleAlaThrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 179
Db 373 ACATGTGCTTTTAAGTAGTCTTATCATCTCAATGGGAACAACCTGGGTATTTGAGAGGTC 432
Qy 180 MetCysArgIleThrValValPheThrGlyValMetCysAlaIleLeuIleLeu 199
Db 433 CTGTGCGGGCCACCACTGATCTTCTATGGCAACATGATGCTCTCATCTTCTGCTCT 492
Qy 200 ThrCysMetGlyIleAsnArgTrpLeuAlaThrAlaHisProPheThrTrpGlnLysLeu 219
Db 493 GCCTGCATCAGCATCAACCGCTACCTGGCCATCTGCTCATCTTTCACCTACCGGGCGTG 552
Qy 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTrp 239
Db 553 CCCAAGCACACTTATGCTTGGTAACATGATGGAGTGTGGGCAACAGTTTCTTATAT 612
Qy 240 MetLeuProPheValIleLeuLysGlnGlyThrHisLeuValHisSerGluIleThr 259
Db 613 ATGCTGCCATTTTCATCTAGCAGCAGGATATATCTTGTTCAGCCAGACATCACACC 672
Qy 260 CysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTrpPheVal 279
Db 673 TGCCATGATGTTCAACAACACTTGGAGTCTCATCTCCCTTCCAACTCTATTACTTCATC 732
Qy 280 SerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTrpThr 299
Db 733 TCCTTGGCATCTTTTGGATCTTAAATTCATTTGTGCTTATCATCTACTGCTATGACGC 792
Qy 300 LeuIleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTrpIleLysAlaValLeu 319
Db 793 ATCATCGGACACTTAATGATACATACATAGATGTTGTTGTTGTTGTTGTTGTTGTTG 852
Qy 320 LeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIleLeuValIleHis 339
Db 853 CTCATCTTGTGATTTTACCATTGCTTGTCTCCAAACAATATTTATTTATTATTATTCAC 912
Qy 340 HisAlaAsnTrpTrpTrpHisAsnThrAspSerLeuTrpPheMetTrpLeuIleAlaLeu 359
Db 913 CATGCTAACTACTACTACCAACAACACTGATGGCTTATATTTATATATCATCATAGCTTTG 972
Qy 360 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTrpPheValMetSerLys 378
Db 973 TGCTGGGTAGTCTTAAATAGTGTGTTAGATCCATCTTATTTTCTTCATGTCATAA 1029
RESULT 4
US-08-742-440A-2
```

```
Sequence 2, Application US/08742440A
Patent No. 5892014
GENERAL INFORMATION:
APPLICANT: Coughlin, Shaun
APPLICANT: Ishihari, Hiroaki
TITLE OF INVENTION: Protease Activated Receptor
TITLE OF INVENTION: 3 and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Avenue, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,440A
FILING DATE: 30-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: UCAL/060PAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1124 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
FEATURE:
US-08-742-440A-2
Alignment Scores:
Pred. No.: 4,48e-126 Length: 1124
Score: 1300.50 Matches: 265
Percent Similarity: 71.99% Conservative: 10
Best Local Similarity: 69.37% Mismatches: 16
Query Match: 60.88% Indels: 92
DB: 2 Gaps: 4
US-09-208-629F-3 (1-407) x US-08-742-440A-2 (1-1124)
Qy 29 PheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSerAlaLysPro 48
Db 15 TTCCTTTCAATACAC-----GGCATAAATGTTTCAGACAACTCAGCAAGCCCA 62
Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 63 ACCTTAACATTATAAGAGTTTAAATGGGGTCCCAAAATACCTTTTGAAGAAATTCNN--- 118
Qy 69 SerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAsp 88
Db 118 ----- 118
Qy 89 SerIleSerThrLeuHisValAsnAlaThrIleGlyTrpLeuArgSerSerLeuSer 108
Db 119 NNNNTACAACTCTCCATGTGTAATAATGTTACCTGGGATACCTGAGAGTTCTTAAGT 178
Qy 109 ThrGlnValIleProAlaIleTrpIleLeuPheValGlyValProSerAsnIle 128
```



```
Db 436 GCCGACCTCTCTCTGTCATCTGGTCCCTCCCTGAAGATCTCCACCACTACATGGCAAC 495
Qy 173 AsnTrrpValPheGlyGluValMetCysArgileThrThrValValPheTyrGlyAsnMet 192
Db 496 AACTGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATG 555
Qy 193 TrrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
Db 556 TATTGCTCCATCTCTTTCATGACCTGCTCAGCGTCAGAGGTACTGGGTGATCGTGAAC 615
Qy 213 ProPheThrTyrGlnTyrLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Db 616 CCATGGGACACCCAGG--AAGAAGGCAAAATCCCGTGGCTCTCTTGGCAATC 672
Qy 233 TrpValMetValPheLeuTyrMetPheValIleLeuLysGlnGluTyrHisLeu 252
Db 673 TGGCTCTGATTTTCTGGTCACCATCTTGTATGTCATGAAGCAGACCATCTACATT 732
Qy 253 ValHisSerGluIleThrThrCysHisAspValValAspLysCysGluSerProSerSer 272
Db 733 CCAGCATTTGAACATCACACCTGTCCAGCATGTGCTGCTGAG--GAGGTATTGGTGGG 789
Qy 273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheIleLeuProPheValIle 292
Db 790 GACATGTTCAATTAATCTCTCTACTGCGCAATGGAGTCTTCCTGTTCCCGGCCCTCCT 849
Qy 293 IleIlePheCysTyrThrThrLeuIle-----His 302
Db 850 ACTGCATCTGCTACGTGCTCATGATCAAGCGTCCGCTCTTCTGCTGATGGATGAACAC 909
Qy 303 LysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeu 322
Db 910 TCAGAGAAGAAAGGAGAGGCT-----ATCCGACTCATCATCACCGTCTG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db 958 GCCATGACTTACTCTGCTTGTCTCTAGCAACCTCTGCTGCTGATGTCATTATTTCCTA 1017
Qy 343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db 1018 ATCAAAACCCAGGAGAGCAGCTCTACGCCCTTACCTTGTGCGCCCTCTGCTCTG 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1078 ACCCTCAACAGTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 1125
```

RESULT 7

```
US-08-476-976-60
; Sequence 60, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2732 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..1269
; US-08-476-976-60

Alignment Scores:
Pred. No.: 5,21e-51 Length: 2732
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservative: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 2 Gaps: 10

US-09-208-629f-3 (1-407) x US-08-476-976-60 (1-2732)
Qy 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuLeuValAlaAlaGlyLeu 27
Db 61 CCAGTCCGGGATGGGAAGTCTCAGCCTGGCTGCTGCTG-----GGAGGTATC 111
Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
Db 112 ACCCTTCTGGCGCCTCGGTCTCTCTCAGCGGACCGAGAACCTTGCCCGGACGCAAC 171
Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
Db 172 AACAGTAAAGGAAGAAAGTCTTATTGCGAGATTAGAAACCCAGCCTCAAATCACTGGGAAA 231
Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
Db 232 GGGGTTCCGGTAGAACACGAGCTTTTCCATCGATGAGTCTCTCGCTCCATCTCACCGGG 291
Qy 74 TrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
Db 291 ----- 291
Qy 94 HisValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
Db 292 -----AAGCTGACCAACGGTCTTTCTTCG 315
Qy 114 AlaIleTyrIleLeuLeuPheValGlyValProSerAsnIleValThrLeuTrpLys 133
Db 316 GTCGCTACATTAATGTTGTTGATGTTTCCAGTAATGCGATGCGCCCTCTGATC 375
Qy 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
Db 376 TTCCTTTTCCGAACGAAGAACAACACCCCGCGTGTATTACATGGCAACCTGGCCTG 435
Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
Db 436 GCCGACCTCTCTCTCTCTCTGTCATCTGTTCCCTCGAAGATCTCTACCACCTACATGGCAAC 495
Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
Db 496 AACTGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATG 555
Qy 193 TyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
```

556 TATTGCTCCATCTCTTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615
213 ProPheThrTyrGlnIleValProLysLeuMetCysSerPheSerLeuMetCysGlyIleVal 232
616 CCCATGGGACACCCAGG---AAGAGGCAAACTCCGCTGGCGTCTCTTGGCAATC 672
233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeu 252
673 TGGCTCTGATTTTCTGTCACCATCTTGTATGTCATGAGCAGACCATCTACAT 732
253 ValHisSerGluIleThrCysHisAspValValAspAlaCysGluSerProSerSer 272
733 CCAGCATTAACATCACCATCTGTCACCATCTTGTATGTCATGAGCAGACCATCTACAT 789
273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
790 GACATGTTCAATTAATCTCTCTACTGCGCATTTGGAGTCTTCTGTTCCCGGCCCTCTT 849
293 IleIlePheCysTyrThrThrLeuIle-----His 302
850 ACTGCATCTGCTGCTGCTCATGATCAAGACGCTCGCTCTCTCTGCTATGGATGAACAC 909
303 LysLeuLysSerTyrAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeu 322
910 TCAGAGAAGAAAGGAGGAGGCT-----ATCCGACTCATCATCACCGTGTG 957
323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHisAlaAsn 342
958 GCCATGACTTCACTCTCTTGTCTCTAGCAACCTCTGCTGCTAGTGCATTATTCCTA 1017
343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
1018 ATCAAAACCCAGGAGGAGGAGGCTCTAGCCCTCTACCTTGTGCGCTCTGCTGCTG 1077
363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerTyr 378
1078 ACCCTCAACAGCTCATAGACCCCTTGTCTATTACTTTGTCTCAAAA 1125

RESULT 8

US-08-474-410-60
Sequence 60, Application US/08474410
Patent No. 6043212
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,410
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 2732 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 73..1269
US-08-474-410-60
Alignment Scores:
Pred. No.: 5,21e-51 Length: 2732
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservative: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 3 Gaps: 10
US-09-208-629F-3 (1-407) x US-08-474-410-60 (1-2732)
QY 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuIleValAlaAlaGlyLeu 27
DB 61 CCACGTCGGGGATCGAAGTCTCAGCGTGGCGTCTGCTG-----GGAGGTATC 111
QY 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
DB 112 ACCCTTCTGCGCGCTCGTCTCTGCGCGGACCGAGAACCTTCACCGGACGCAAC 171
QY 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
DB 172 AACAGTAAGGAGAGAGTCTTATGGCAGATTAGAACCCAGCTCCAATCATCTGGAAA 231
QY 54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
DB 232 GGGGTTCCGGTAGAACCGAGCTTTTCCATCGATGAGTTCTCTGCGTCCATCTCCCGGG 291
QY 74 TrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
DB 291 ----- 291
QY 94 HisValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
DB 292 -----AAGCTGACCAACGGTCTTCTTCG 315
QY 114 AlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLys 133
DB 316 GTCGCTCTACATTATTGTTGTTGATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 375
QY 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
DB 376 TTCCTTTTCCGAACGAGAGAAACACCCCGCGTGTATTATCATGGCCAACTGGCGCTT 435
QY 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
DB 436 GCGGACCTCTCTCTCTCATCTGGTTCCCGCTGAGAGATCTCTTACCACCTACATGGAAC 495
QY 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
DB 496 AACTGGGTCTACGGGAGGCGCCCTGTCAAGGTCTCATTTGGCTTTTCTATGTTAATC 555
QY 193 TyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
DB 556 TATTGCTCATCTCTTCTCATGACCTGCTGAGGTGCTGAGGTGCTGAGTCTGAGT 615
QY 213 ProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
DB 616 CCATGGGACACCCAGG---AAGAGGCAAACTCCGCTGGCGTCTCTTGGCAATC 672
QY 233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeu 252

```
Db 673 TGGCTCTGATTTTCGGTACATCCCTTTGTATGTCATGAAGCAGACCATCTACATT 732
Qy 253 ValHisSerGluIleThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Db 733 CCAGCATTTGAACATACACCATCTGTCAGATGTCGCTGAG--GAGGTATTGGTGGG 789
Qy 273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Db 790 GACATGTTCAATTAATCTCTCTCCTCAGTGGCCTGAGTCTCTCTGTTCCGGGCTCTCT 849
Qy 293 IleIlePheCysTyrThrThrLeuIle-----His 302
Db 850 ACTGCATCTCGCTACGTCTCATGATCAAGACGCTCGCTCTCTCTGATGATGAACAC 909
Qy 303 LysLeuLysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuLeu 322
Db 910 TCAGAGAGAAAGGAGAGAGGCT-----ATCGACTCATCATCAGCGGTG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db 958 GCCATGTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1017
Qy 343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db 1018 ATCAAAACCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1078 ACCCTCAACAGCTGCATAGACCCCTTGTCTATTAATCTCTCTCTCTCTCTCTCT 1125
```

RESULT 9

```
US-08-486-673B-60
; Sequence 60, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486.673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 60
; LENGTH: 2732
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (73)..(1269)
; OTHER INFORMATION: C140 receptor, cDNA and deduced protein sequences
US-08-486-673B-60
```

Alignment Scores:

Pred. No.:	5,21e-51	Length:	2732
Score:	584.00	Matches:	128
Percent Similarity:	51.52%	Conservative:	76
Best Local Similarity:	32.32%	Mismatches:	126
Query Match:	27.34%	Indels:	66
DB:	4	Gaps:	10

US-09-208-629F-3 (1-407) x US-08-486-673B-60 (1-2732)

```
Qy 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuValAlaAlaGlyLeu 27
Db 61 CCAGTCCGGGATGCGAAGTCTCAGCTGCGGTGCTGCTG-----GGAGGTATC 111
Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
```

```
Db 112 ACCCTTCTGGCGGCTCTCGTCTCTCTGACCGGACCGAAGAACTTCACCGGACGCAAC 171
Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
Db 172 AACAGTAAAGGAAGAGTCTTATTGGCAGATTAGAAACCCAGGCTCCAATCACTCGGAAA 231
Qy 54 SerPheAsnGlyCysProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
Db 232 GGGGTTCCGGTAGAACACGAGCTTTTCCATCGATGAGTCTCTCGCGTCCATCTCACCAGG 291
Qy 74 TrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
Db 291 ----- 291
Qy 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
Db 292 -----AAGCTGACCAAGGCTCTTCTTCCG 315
Qy 114 AlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLys 133
Db 316 GTCGCTACATATTGTGTTGTGATTGTTGCCAGTAATGGCATGGCCCTCTGATC 375
Qy 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
Db 376 TTCCTTTTCCGAACGAAGAAAGAACACCCCGCGTGATTTACATGGCCCAACCTGGCGCTG 435
Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
Db 436 GCGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 495
Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
Db 496 AACTGGGTCTACGGGAGGCGCTGTCAAGGTCTCATTTGGCTTTTCTATGGTAACATG 555
Qy 193 TyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrIleuAlaThrAlaHis 212
Db 556 TATTGCTCCATCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 615
Qy 213 ProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Db 616 CCCATGGGACACCCCGAGG---AAGAAGGCAACATCGCGTGGCGTCTCTCTGGCNAATC 672
Qy 233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeu 252
Db 673 TGGCTCTCTGATTTTCTGGTCAACATCCCTTTGTATGTCATGAAGCAGACCATCTACATT 732
Qy 253 ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Db 733 CCAGCATTTGAACATCACCATCTGTCAGATGTCGCTGCTGCTGAG---GAGGTATTGGTGGG 789
Qy 273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Db 790 GACATGTTCAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 849
Qy 293 IleIlePheCysTyrThrThrLeuIle-----His 302
Db 850 ACTGCATCTCGCTACGTCTCATGATCAAGACGCTCGCTCTCTCTGATGATGAACAC 909
Qy 303 LysLeuLysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuLeu 322
Db 910 TCAGAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db 958 GCCATGTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1017
Qy 343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db 1018 ATCAAAACCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1078 ACCCTCAACAGCTGCATAGACCCCTTGTCTATTAATCTCTCTCTCTCTCTCTCTCT 1125
```


APPLICATION NUMBER: US 08/390,301
 FILING DATE: 25-JAN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: ADLER, REID G.
 REGISTRATION NUMBER: 30,988
 REFERENCE/DOCKET NUMBER: 2803-0006.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 TELEX: 90-4030
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1475 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 232..1416
 NAME/KEY: mat_peptide
 LOCATION: 232
 US-08-476-000-1

Alignment Scores:

Pred. No.: 6,25e-51 Length: 1475
 Score: 579.50 Matches: 111
 Percent Similarity: 60.00% Conservatives: 66
 Best Local Similarity: 37.63% Mismatches: 101
 Query Match: 27.13% Indels: 17
 DB: 1 Gaps: 5

US-09-208-629F-3 (1-407) x US-08-476-000-1 (1-1475)

Qy 95 ValAsnAlaThrIleGlyTyrLeuArgSerLeuSerThrGlnValIleProAla 114
 Db 406 ATGATGAGTCTCTGGCGTCATCTCACCGGAGCTGACCGGTCTTCTCCGGTC 465
 Qy 115 IleTyrIleLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
 Db 466 GTCTACATTATTGTGGTTGATTGGTTGCCAGTAATGGCATGGCCCTCTGGATCTC 525
 Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
 Db 526 CTTTTCGACGAGAGAAACACCCCGCGCTGATTACATGGCCAACTGGCGCTGGCC 585
 Qy 154 AspLeuLeuPheCysValThrLeuProPheIleAlaTyrHisLeuAsnGlyAsnAsn 173
 Db 586 GACCTCTCTGTGTCATCTGGTCCCTGAGAGATCTCTACCACTACATGGCAACAC 645
 Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
 Db 646 TGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATGTAT 705
 Qy 194 CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
 Db 706 TGCTCCATCTCTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCC 765
 Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrp 233
 Db 766 ATGGGACACCCAGG---AAGAAGGCAACATCCCGTGGCGTCTCTTGGCAATCTGG 822
 Qy 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
 Db 823 CTCCTGATTTTCTGGTCACCATCTTGTATGTGTCATGAGCAGACCATCTACATCCA 882
 Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
 Db 883 GCATTGAACATCACCATCTGTCACCATGTGCTGCTGAG---GAGGTATTGGTGGGGAC 939
 Qy 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
 Db 940 ATGTTCAATTACTTCTCTACTGGCCATTGGAGTCTTCTGTTCCTCCGCGCCCTCTTACT 999

Qy 294 IlePheCysTyrThrThrLeuIleHisLysLeuLysSer----- 306
 Db 1000 GCATCTGCCTAGCTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGATGAACACTCA 1059
 Qy 307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleVal 323
 Db 1060 GAGAACAAAGGAGAGGCT-----ATCCGACTCATCATCACCGGTGGGCC 1107
 Qy 324 IlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyr 343
 Db 1108 ATGTACTTCTGCTTGTCTCTAGCAACCTCTCTGCTAGTGCATTATTTCCCTAATC 1167
 Qy 344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer 363
 Db 1168 AAAACCCAGAGGAGGACGCTACGGCCTCTACCTTGTGCGCCCTCTGCTGTGCGACC 1227
 Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
 Db 1228 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 1272

RESULT 12

US-08-472-840-1
 ; Sequence 1, Application US/08472840
 ; Patent No. 5763575
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FORSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/472,840
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1475 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 232..1416
 ; NAME/KEY: mat_peptide
 ; LOCATION: 232
 ; US-08-472-840-1

Alignment Scores:


```
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db 646 TGGGTCTACGGGAGGCGCTGTGCAAGGTCTCATTTGGCTTTTCTATGTAACATGTAT 705
Qy 194 CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGTCTCATCTCTTCATGACTGCTCAGCGTGAGAGGTACTGGGTGATCGTGAACCCC 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
Db 766 ATGGGACACCCAGG---AAGAAGGCAACATCGCGTGTGGGCTCTCTTGGCAATCTGG 822
Qy 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
Db 823 CTCCTGATTTTCTGGTCAACCTCTGTGATGATGATGATGATGATGATGATGATGATGAT 882
Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
Db 883 GCATTGAACATCACCATCTGTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
Qy 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
Db 940 ATGTTCAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 999
Qy 294 IlePheCysTyrThrThrLeuIleHisLysLeuLysSer----- 306
Db 1000 GCATCTGCCCTACGCTCATGATCAAGACGCTCGCTCTCTCTCTCTCTCTCTCTCTCT 1059
Qy 307 -----LysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuLeuVal 323
Db 1060 GAGAACAAAGGAGAGGGGT-----ATCCGATCATCATCCGCTGCTGCTGCC 1107
Qy 324 IlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyr 343
Db 1108 ATGTAATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1167
Qy 344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySer 363
Db 1168 AAAACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1227
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1228 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 1272

RESULT 14
US-08-474-410-1
; Sequence 1, Application US/08474410
; Patent No. 5043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 232..1416
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 232
; US-08-474-410-1
Alignment Scores:
Pred. No.: 6,25e-51 Length: 1475
Score: 579.50 Matches: 111
Percent Similarity: 60.00% Conservative: 66
Best Local Similarity: 37.63% Mismatches: 101
Query Match: 27.13% Indels: 17
DB: 3 Gaps: 5
US-09-208-629f-3 (1-407) x US-08-474-410-1 (1-1475)
Qy 95 ValAsnAlaThrIleGlyTyrLeuArgSerSerSerSerThrGlnValIleProAla 114
Db 406 ATCATGAGTCTCTCGCTCCATCTCACGGGAAGCTGACACGGTCTTCTTCGGGTC 465
Qy 115 IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
Db 466 GTCTACATTATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 525
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 526 CTTTTCGGAACGAAGAAGAAACACCCCGCGTGATTATCATATGGCCCAACCTTGGGCT 585
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 586 GACCTCTCTCTCATCTCTGTTCCCTCCCTGAAGATCTCTACACCTACATGCGCAACAC 645
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db 646 TGGGTCTACGGGAGGCGCTGTGCAAGGTCTCATTTGGGCTTTTCTATGTAACATGTAT 705
Qy 194 CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGTCTCATCTCTTCATGACTGCTCAGCGTGAGAGGTACTGGGTGATCGTGAACCCC 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
Db 766 ATGGGACACCCAGG---AAGAAGGCAACATCGCGTGTGGGCTCTCTTGGCAATCTGG 822
Qy 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
Db 823 CTCCTGATTTTCTGGTCAACCTCTGTGATGATGATGATGATGATGATGATGATGATGAT 882
Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
Db 883 GCATTGAACATCACCATCTGTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
Qy 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
Db 940 ATGTTCAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 999
Qy 294 IlePheCysTyrThrThrLeuIleHisLysLeuLysSer----- 306
```

Db 1000 GCATCTGCCTACGGTCTCATGATCAAGACGCTCGCTCTCTGCTATGGATGAACACTCA 1059
 QY 307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323
 Db 1060 GAGAACAAAGGAGGAGGCT-----ATCGACTCATCATCACCCTGCTGGCC 1107
 QY 324 IlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyr 343
 Db 1108 ATGTACTTCTCTGCTTGTCTCTAGCACTCTGCTGCTAGTGATATTTCCTAATC 1167
 QY 344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySer 363
 Db 1168 AAAACCCAGAGGAGGAGGAGGCTCTACGCTCTCTACCTTGTGCGCTCTGCTGCGACC 1227
 QY 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
 Db 1228 CTCACAGCTGCATAGACCCCTTGTCTATTACTTTGTCTCAAAA 1272

RESULT 15

US-08-486-673B-1
 ; Sequence 1, Application US/08486673B
 ; Patent No. 6297026
 ; GENERAL INFORMATION:
 ; APPLICANT: Sundelin, Johan
 ; APPLICANT: Scarborough, Robert M.
 ; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
 ; FILE REFERENCE: 44481-5006-08-US
 ; CURRENT APPLICATION NUMBER: US/08/486,673B
 ; CURRENT FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: US 08/097,938
 ; PRIOR FILING DATE: 1993-07-26
 ; PRIOR APPLICATION NUMBER: PCT/US94/08536
 ; PRIOR FILING DATE: 1994-07-26
 ; NUMBER OF SEQ ID NOS: 63
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1475
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (232)..(1416)
 ; OTHER INFORMATION: C140 receptor, genomic DNA and deduced protein
 ; OTHER INFORMATION: sequences
 US-08-486-673B-1

Alignment Scores:

Pred. No.: 6.25e-51 Length: 1475
 Score: 579.50 Matches: 111
 Percent Similarity: 60.00% Conservative: 66
 Best Local Similarity: 37.63% Mismatches: 101
 Query Match: 27.13% Indels: 17
 DB: 4 Gaps: 5

US-09-208-629F-3 (1-407) x US-08-486-673B-1 (1-1475)

QY 95 ValAsnAsnAlaThrIleGlyTyrLeuArgSerLeuSerThrGlnValIleProAla 114
 Db 406 ATCGATGAGTTCTCGCTCCATCTCCCGGGAAGCTGACCAGGCTCTTTCTCGGTC 465
 QY 115 IleTyrIleLeuLeuValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
 Db 466 GTCTACATATTGTGTGTGATTGTGCTCCAGTAATGCGATGCGCCCTCTGATCTTC 525
 QY 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
 Db 526 CTTTTCGACGAAGAAGAACACCCCGCGCTGATTTACATGGCCACCTGGCTTGGCC 585
 QY 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
 Db 586 GACCTCTCTCTGTCATCTGTTGTTCCCTTGAAGATCTCTACCACTACATGACCAAC 645

QY 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
 Db 646 TGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATGAT 705
 QY 194 CysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
 Db 706 TGCTCCATCTCTTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAACCCC 765
 QY 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
 Db 766 ATGGGACACCCAGG---AAGAAGCAACATCGCGTTGGCGTCTCTCTGGCAATCTGG 822
 QY 234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
 Db 823 CTCCTGATTTTCTGTCACCATCCCTTTGTATGTCATGACGACGACCATCTACATTTCCA 882
 QY 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
 Db 883 GCATTGAACATCACCACCTGTCCAGATGTCTCCCTGAG---GAGGTATTGGTGGGGAC 939
 QY 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIle 293
 Db 940 ATGTTCAATTACTTCTCTCACTGGCCATTGGAGTCTTCTGTCCCGGCCCTCTTACT 999
 QY 294 IlePheCysTyrThrThrLeuIleHisLysLeuLysSer----- 306
 Db 1000 GCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGATGAACACTCA 1059
 QY 307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323
 Db 1060 GAGAACAAAGGAGGAGGCT-----ATCGACTCATCATCACCCTGCTGGCC 1107
 QY 324 IlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyr 343
 Db 1108 ATGTACTTCTCTGCTTGTCTCTAGCACTCTGCTGCTAGTGATATTTCCTAATC 1167
 QY 344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer 363
 Db 1168 AAAACCCAGAGGAGGAGGAGGCTCTACGCTCTCTACCTTGTGCGCTCTGCTGCGACC 1227
 QY 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
 Db 1228 CTCACAGCTGCATAGACCCCTTGTCTATTACTTTGTCTCAAAA 1272

RESULT 16

US-08-097-938-3
 ; Sequence 3, Application US/08097938
 ; Patent No. 5629174
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
 ; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
 ; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington, D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/097,938
 ; FILING DATE: 26-JUL-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 22803-20006.00

TELECOMMUNICATION INFORMATION:

	TELEPHONE: (202) 887-1500	
	TELEFAX: (202) 887-0763	
	TELEX: 90-4030	
/	INFORMATION FOR SEQ ID NO: 3:	
/	SEQUENCE CHARACTERISTICS:	
/	LENGTH: 1255 base pairs	
/	TYPE: nucleic acid	
/	STRANDEDNESS: single	
/	TOPOLOGY: linear	
/	FEATURE:	
/	NAME/KEY: CDS	
/	LOCATION: 56..1249	
/	FEATURE:	
/	NAME/KEY: mat_peptide	
/	LOCATION: 56	
/	US-08-097-938-3	

Alignment Scores:		
Pred. No.:	2,57e-49	Length: 1255
Score:	563.00	Matches: 115
Percent Similarity:	54.52%	Conservative: 72
Best Local Similarity:	33.53%	Mismatches: 128
Query Match:	26.36%	Indels: 28
DB:	1	Gaps: 7

US-09-208-629F-3 (1-407) x US-08-097-938-3 (1-1255)

QY	49 ThrLeuThrIleLysSerPheAsnGlyClyProGlnAsnThrPheGluGluPheProLeu 68	
Db	116 ACCCTTGTCCTTCTTTGTACAGGAACCAATAGATCCTCTAAAGGAAGCCCTATT 175	
QY	69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrIleLysAlaGluCys 85	
Db	176 GGTAAAGGTGTGGGCACATCCACGTCACTGGAAGAGGATTACAGTTGAA----- 226	
QY	86 ProGluAaspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105	
Db	227 -----ACAGTCTTTTCTGTGGATGATTTTCTGCATCTGCTCCTACTGGA 271	
QY	106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyWalPro 125	
Db	272 AAACATGACCACTGTCTTCCTTCCAATTGCTCACACAATTGTGTTGTGGTGGTGGTGGCA 331	
QY	126 SerAsnIleValThrLeuTriplysLeuSerLeuargThr---LysSerIleSerLeuVal 144	
Db	332 ACTAACGGCATCGGCCCTGTGGTCTTCTTTCCGAATAAGAAGACACCCTGCTGTG 391	
QY	145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164	
Db	392 ATTTACATGGCCAATCTGGCCTTGGCTGACCTCTCTCTCATCTGGTTCCTCTTGAAG 451	
QY	165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184	
Db	452 ATTGGCTATCACATACATGCGCAACAACATGGATTATGGGGAAGCTCTTGTATATGTGCT 511	
QY	185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204	
Db	512 ATTGGCTTTTCTATGGCAACATGACTGTGTTCATTCTTTCATGACCTGCTCCTCAGTGG 571	
QY	205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224	
Db	572 CAGAGGTATGGGTCTATCGTAACCCCATGGGCACCTCCAG--AAGAAGGCCAAACATT 628	
QY	225 SerLeuLeuMetCysGlyIleValITrpValMetValPheLeuTyrMetLeuPropheVal 244	
Db	629 GCATTTGGCATCTCCCTGGCAATATGGCTGTGTAATCTGTGGTGCACCATCCCTTTGTAT 688	
QY	245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisaspValVal 264	
Db	689 GTTCGTGAAGCAGCACCATCTTCATCTCCTGCCCTGAACATCACGACCTGTCTATGATGTTTTG 748	
QY	265 AspAlaCysGluSerProSerSerPhe-----ArqPheTyrTyrPheValSer 280	

US-09-208-629F-3 (1-407) x US-08-472-840-3 (1-1255)

Qy	49	ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu	68
Db	116	: : : : :	175
Qy	69	SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys	85
Db	176	: : : : : : : : : :	226
Qy	86	ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer	105
Db	227	: : : : : : : : : : : :	271
Qy	106	SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro	125
Db	272	AACTGGACCACTGCTTCCTCCCAATTGCTACAAATTGGTTGTGGTGGTTCGCCA	331
Qy	126	SerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSerLeuVal	144
Db	332	: : :	391
Qy	145	IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys	164
Db	392	ATTTACATGGCCAATCTGGCCCTCTGGCTGACCTCTCTCTGTCATCTGGTTCCTTCGAAG	451
Qy	165	IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr	184
Db	452	ATTGCCTATCACATACATACGCAACAACCTGGATTTATGGGAAGCTCTTTGTAATGTGCTT	511
Qy	185	ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyLe	204
Db	512	ATTGGCTTTTTCATGGCAACATGACTACTGTGTCCATTCTCTTCATGACCTGCCTCAGTGTG	571
Qy	205	AsnArgTyrLeuAlaThrAlaHisSerProPheThrTyrGlnLysLeuProLysArgSerPhe	224
Db	572	CAGAGGTATTGGGTATCGTGAACCCCATGGGGCATTCCAGG---AAGAAGGCAACAATT	628
Qy	225	SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal	244
Db	629	GCCATTTGGCATCTCCCTGGCAATATGGCTGTGATTTCTGTGGTCACCACTCCCTTGTAT	688
Qy	245	IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal	264
Db	689	GTCGTGAAGCAGACCATCTTCATTCCTGCCCTGAAACATCACGACCTGTCATGATGTTTG	748
Qy	265	AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer	280
Db	749		793
Qy	281	LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu	300
Db	794	CTGGCCATTGGGGTCTTTCTGTTCCAGCCTTCTCACAGCCTCTGCTATGCTGCTGATG	853
Qy	301	IleHisLysLeuLysSer-----LysAspArgIleIleTrpLeuGlyTyr	314
Db	854	ATCAGAAATGCTGCGATCTCTGCCATGGATGAAAACCTCAGAGAAGAAAAGGAGGGCC	913
Qy	315	IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle	334
Db	914	ATCAAACTCATGTGTCACCTCTCGGCCATGTACCTGATCTGCTTCACTCCTAGTAACCTT	973
Qy	335	IleLeuValIleHisHisAlaAsnTyrTyrThrHisAsnThrAspSerLeuTyrPheMet	354
Db	974	CTGCTTGGTGGCATTAATTTCTGATTAAAGCCAGGCCAGACCCATGCTATGCCCCCTG	1033
Qy	355	TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe	374
Db	1034	TACATTGAGCCCTCTGCCCTCTCTACCCCTTAACAGCTGCATCGACCCCTTGTCTATTAC	1093
Qy	375	ValMetSer	377
Db	1094	TTTGTTCATA	1102

RESULT 19

```

US-08-476-976-3
; Sequence 3, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..1249
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 56
; US-08-476-976-3

Alignment Scores:
Pred. No.: 2,57e-49 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 2 Gaps: 7

US-09-208-629F-3 (1-407) x US-08-476-976-3 (1-1255)

Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
|||||
Db 116 ACCCTTGCTCTCTTCTTGTCAGGAACCATAGATCTCTTAAGGAGAGACCTTATT 175
|||||

Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
|||||

Db 176 GGTAAGTTTCATGGCACCCACGTCCTCGAAAGGAGGTTACACTTGA----- 226
|||||

Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
|||||

Db 227 -----ACAGTCTTTTCTGTGGATGAGTTTCTGTCATCTGTCTCTCACTGGA 271
|||||

```

```
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 272 AAACAGCACTGCTCTCCCTCCCAATGCTACACAATGGTTGGTGGGGTTGCCA 331
QY 126 SerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
DB 332 AGTAACGGCATGGCCCTGGTGGTCTTTCCGAACTAAGAAGAACAGCACCCTGCTGTG 391
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspIleLeuPheCysValThrLeuProPheLys 164
DB 392 ATTACATGCGCAATCTGGCTGGCTGACCTCTCTCTCATCTGGTTCCCTTGAAG 451
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCysArgIleThr 184
DB 452 ATTCGCTATACATACATGCAACCACTGGATTATGGGAAGCTCTTTGTAATGCTT 511
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIle 204
DB 512 ATTGGCTTTTCTATGCAACATGCTACTGTTCCATTCTCTCATGACCTGCCCTCAGTGTG 571
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
DB 572 CAGAGGTAATTGGGTCATCGTCAACCCCATGGGCACCTCCAGG---AAGAAGGCAACATT 628
QY 225 SerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeuProPheVal 244
DB 629 GCCATTGGCATCTCCCGCAATATGCTGCTGATCTCTGCTGTCACCATCCCTTGTAT 688
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
DB 689 GTCGTGAAGCAGACCATCTTCATTCCTGCTCCCTGCAACATCAGACCTGTCATGTTTG 748
QY 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
DB 749 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCT 793
QY 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
DB 794 CTGGCCATTGGGCTCTTCTGTTCCAGCCTCTCTCACGCCCTGCTGCTATGCTGTGAT 853
QY 301 IleHisLysLeuLysSer-----LysAspArgIleTyrLeuGlyTyr 314
DB 854 ATCAGATGCTGCGATCTCTGCTCCATGATGAACTCAGAGAAGAAAGAGAGAGGCC 913
QY 315 IleLysAlaValLeuLeuIleValIlePheThrIleCysPheAlaProThrAsnIle 334
DB 914 ATCAAACTCATTTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973
QY 335 IleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
DB 974 CTGTTGTTGGTCAATATTTCTGATTAAAGCCAGGCCAGCCCATGCTATGCCCTG 1033
QY 355 TyrLeuIleAlaCysLeuGlySerLeuAsnSerCysLeuAspPropheLeuTyrPhe 374
DB 1034 TACATTGTAGCCCTCTGCTCTCTACCTTAAAGCTGATGCTGACCCCTTTGCTATTAC 1093
QY 375 ValMetSer 377
DB 1094 TTTGTTTCA 1102
RESULT 20
US-08-474-410-3
; Sequence 3, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
```

```
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,410
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 56..1249
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 56
US-08-474-410-3
Alignment Scores:
Pred. No.: 2,57e-49 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 7 Gaps: 7
US-09-208-629F-3 (1-407) x US-08-474-410-3 (1-1255)
```

```
QY 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
DB 116 ACCCTTGCTCTCTCTTCTTGTACAGGACCAATAGATCTCTTAAGAGGAAGACCTTATT 175
QY 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
DB 176 GGTAAGGTTTGGTGGCATCCACAGCTCCTGGGAAAGAGGATTACAGTTGAA----- 226
QY 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
DB 227 -----ACAGTCTTTTCTGTGGATGAGTTTCTGTCATCTGTCTCCTCCTGGA 271
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 272 AAACAGCACTGCTCTCTCTTCCATGCTACACAATGCTGTTGGTGGGGTTGCCA 331
QY 126 SerAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
DB 332 AGTAACGGCATGGCCCTGGTGGTCTTTCTTTCCGAACTAAGAAGAACAGCACCCTGCTGTG 391
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
DB 392 ATTACATGCGCAATCTGGCTGGCTGACCTCTCTCTGCTGCTGCTGCTGCTGCTGGAAG 451
```


QY 165 IleAlaTyrHisLeuAenGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 452 ATTGGCTATCACATACATGCAACAACTGGATTATGGGAAGCTCTTTGTAATGCTT 511
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
Db 512 ATTGGCTTTTCTATGCAACATGCTGTTCCATTCTCTCATGACCTGCCATGCTG 571
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 572 CAGAGGTATTGGGTTCATCGTGAACCCATCGGGCACTCCAGG--AAGAAGCAAAACATT 628
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 629 GCCATTGGCATCTCCCTGGCAATATGCTGATCTCTGCTGGTGCACCATCCCTTTGAT 688
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
Db 689 GTCGTGAAGCAGACCATCTTCAITTCCTGCCCTGACATCAGACCTGCTCATGATGTTT 748
QY 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
Db 749 -----CCTGAGCAGCTCTTGTGGGAGACATGTTCAATTACTTCTCTCT 793
QY 281 LeuAlaPhePheGlyPheLeuIleValTrpValMetValPheLeuTyrMetLeuProPheVal 300
Db 794 CTGGCAATGGGTCTTCTGTTCCAGCCTTCTCACAGCCTCTGCTCATGCTGCTGATG 853
QY 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 854 ATCAGAAATGCTGGATCTTCTGCTGAGTGAACAACTCAGAGAAGAAAGAGAGGCC 913
QY 315 IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 914 ATCAAACTCATGTACATGCTCTGGCCATGTACCTGATCTGCTTCATCTCTAGTAACCT 973
QY 335 IleLeuValIleHisHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
Db 974 CTGCTTGTGGTGAATATTCTTCTGATTAAAGAGCCAGGCCAGCATGCTATGCCCTG 1033
QY 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
Db 1034 TACATTGAGCCCTCTGCTCTCTACCCCTTAACAGCTGCATCGACCCCTTTGCTATTAC 1093
QY 375 ValMetSer 377
Db 1094 TTTGTTTCA 1102

RESULT 21

US-08-486-673B-3
; Sequence 3, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; APPLICANT: Scarborough, Robert M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486.673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)..(1249)
; OTHER INFORMATION: C140 receptor, genomic DNA and deduced protein

; OTHER INFORMATION: sequences
US-08-486-673B-3

Alignment Scores:
Pred. No.: 2 57e-49 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 4 Gaps: 7

US-09-208-629f-3 (1-407) x US-08-486-673B-3 (1-1255)

QY 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 116 ACCCTTGTCTTCTTCTTGTACAGGAACCAATAGATCTCTAAAGGAAGACCTTATT 175
QY 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
Db 176 GGTAAAGTTGATGGGCACATCCCACTGCTGGAAGAGGATTACAGTTGAA----- 226
QY 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
Db 227 -----ACAGTCTTTTCTGGATGAGTCTTCTGCATCTGCTCCTCACTGGA 271
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Db 272 AAATCAGCACCTGCTTCTTCCCAATGTCTACACAATGTGTTGTGGTGGGTGGCCA 331
QY 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
Db 332 AGTAACGGCATGCCCTGTGGGTCTTCTTCCGAACAAAGAAAGACGACCCTGCTGTG 391
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 392 ATTATCATGGCAATCTGGCTTGGCTGACCTCTCTCTGTCTATCTGGTTCCTTGAAG 451
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 452 ATTGCCTATCACATACATACGCAACAACTGATTATATGGGAAGCTCTTTGTAATGCTT 511
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
Db 512 ATTGGCTTTTCTATGGCAACATGATGTTCTTCAATCTCTTCATGACCTGCTCAGTGTG 571
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 572 CAGAGGTATTGGGTTCATCGTGAACCCATCGGGCACTCCAGG--AAGAAGCAAAACATT 628
QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 629 GCCATTGGCATCTCCCTGGCAATATGCTGATCTGCTGGTGCACCATCCCTTTGAT 688
QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
Db 689 GTCGTGAAGCAGACCATCTTCAITTCCTGCCCTGACATCAGACCTGCTCATGATGTTT 748
QY 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
Db 749 -----CCTGAGCAGCTCTTGTGGGAGACATGTTCAATTACTTCTCTCT 793
QY 281 LeuAlaPhePheGlyPheLeuIleValTrpValIleIleIlePheCysTyrThrThrLeu 300
Db 794 CTGGCAATGGGTCTTCTGTTCCAGCCTTCTCACAGCCTTCTGCTCATGCTGCTGATG 853
QY 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 854 ATCAGAAATGCTGGATCTTCTGCTGAGTGAACAACTCAGAGAAGAAAGAGAGGCC 913
QY 315 IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 914 ATCAAACTCATGTCTGCTGAGTGAACAACTCAGAGAAGAAAGAGAGGCC 973
QY 914 ATCAAACTCATGTCTGCTGAGTGAACAACTCAGAGAAGAAAGAGAGGCC 973

Qy	335	IleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet	354
Db	974	CTGCTTGGTGCATATTTCTCGATTAAAGACGACGACCCATGTCGCCCTG	1033
Qy	355	TyrLeuIleAlaLeuCysLeuGlySerIleuAsnSerCysLeuAspPropheLeuTyrPhe	374
Db	1034	TACATTAGCCCTCTGCCTCTACCCCTTAACAGTCGATCGACCCCTTTGCTATTAT	1093
Qy	375	ValMetSer	377
Db	1094	TTTGTTCAT	1102

RESULT 22

US-08-476-000-62
; Sequence 62, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 50..1240
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 50
; US-08-476-000-62

```
Db 1073 TTGTCTATTACTTTGTTTCA 1093
||||:||||:||||:||||:
RESULT 23
US-08-472-840-62
; Sequence 62, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 50..1240
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 50
US-08-472-840-62
Alignment Scores:
Pred. No.: 5,41e-48 Length: 1414
Score: 551.00 Matches: 118
Percent Similarity: 51.77% Conservative: 72
Best Local Similarity: 32.15% Mismatches: 139
Query Match: 25.80% Indels: 38
DB: 1 Gaps: 8
US-09-208-629f-3 (1-407) x US-08-472-840-62 (1-1414)
Qy 22 LeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
Db 74 CTGCTGGGGCGCCATCTCTCTAGCAGCCTCTCTCTCTGCGAGTGGCCACCATCCAAAGGA 133
Qy 42 SerAspSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
Db 134 ACCAATAGATCCTCTAAAGGAGAGCCTTATTGGTAAGGTTGATGCG----- 181
Qy 62 ThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIle 81
Db 182 -----ACATCCACGTC-----ACTGGAAGAGAGTTACAGTT 214
Qy 82 LysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGly 101
Db 215 GAA-----ACAGTCTTTCTGTGGATGAGTTTCTGATCT 250
Qy 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 121
Db 251 GTCTCGCTGGAAACCTGACCACTGCTCTTCCCAATTGCTACACAAATGTGTGTTGCG 310
Qy 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSer 140
Db 311 GTGGGTTTGCCAAAGTAACGGCATATGGCTCTTCTTTTCGGAACATAAGAGAAG 370
Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 371 CACCCTGCTGTGATTATACATGGCCAACTGGCCCTTGGCTGACCTCTCTCTGTCATCTG 430
Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 431 TTCCCTTGAAGATTGCTTATCACATACATGGAACCACTGGATTTATGGGAAGCTCTT 490
Qy 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
Db 491 TGTAAATGTGCTTATTGCTTTTCTATCGCAACATGCTACTGTTCCATTCTCTCATGACC 550
Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 551 TGCTCAGTGTGAGAGGATTGGGTATCGTGAACCCATGGGGCCATCCAGG---AAG 607
Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 608 AAGGCAACATTTGCCATTGCGATCTCCCTGGCAATATGGCTGCTGCTGCTGGTCACC 667
Qy 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCys 260
Db 668 ATCCCTTTGTATGTCGTGAAGCAGACCATCTTCTATTCCTGCTGAAACATCAGGACCTGT 727
Qy 261 HisAspValValAspAlaCysGluSerProSerPhe-----ArgPheTyr 276
Db 728 CATGATGTTTGT-----CCTGAGCAGCTCTTGTGGGAGACATGTTCAAT 772
Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCys 296
Db 773 TACTTCTCTCTCTGCGCATTTGGGCTCTTCTGTTCCAGCCTTCTCTCAGCCTCTGCC 832
Qy 297 TyrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
Db 833 TATGTCGTGATGATCAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 892
Qy 311 TrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIlePheAla 330
Db 893 AGGAAGAGGGCCATCAAACTCATTTGCTGCTGCTGCGCATGTACCTGATGTACCTGCTCACT 952
Qy 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
Db 953 CCTAGTAACCTTCTGCTTGTGGTGCAATTTTCTGATTAAAGACCCAGGSCCAGAGCCAT 1012
Qy 351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
Db 1013 GTCTATGCCCTGTACATTTGTAGCCCTCTGCTCTCTACCCCTTAACAGCTGATCGACCCC 1072
Qy 371 PheLeuTyrPheValMetSer 377
Db 1073 TTGTCTATTACTTTGTTTCA 1093
RESULT 24
US-08-476-976-62
; Sequence 62, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
```


Db 1465 ATTACTATTACGCTCC 1482

RESULT 28

US-08-767-993-12

Sequence 12, Application US/08767993

Patent No. 6010885

GENERAL INFORMATION:

APPLICANT: TURNER, George J.

APPLICANT: BETLACH, Mary C.

TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS POLYPEPTIDES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESS: Walter H. Dreger

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/767,993

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Walter H.

REGISTRATION NUMBER: 24,190

REFERENCE/DOCKET NUMBER: A-57669/WH

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1764 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: repeat region

LOCATION: 435..462

OTHER INFORMATION: /note= "Sequence encoding

Patent No. 6010885

OTHER INFORMATION: polypartitic acid."

FEATURE:

NAME/KEY: misc feature

LOCATION: 463..465

OTHER INFORMATION: /note= "Codon encoding the

OTHER INFORMATION: N-terminal amino acid of the human thrombin

OTHER INFORMATION: receptor protein."

FEATURE:

NAME/KEY: misc feature

LOCATION: 1630..1632

OTHER INFORMATION: /note= "Codon encoding the

OTHER INFORMATION: C-terminal amino acid of the human thrombin

OTHER INFORMATION: receptor protein."

FEATURE:

NAME/KEY: repeat region

LOCATION: 1633..1650

OTHER INFORMATION: /note= "Sequence encoding

Patent No. 6010885

OTHER INFORMATION: polyhistidine."

FEATURE:

NAME/KEY: misc feature

LOCATION: 648..656

OTHER INFORMATION: /note= "Deleted AlwNI restriction

OTHER INFORMATION: site."

FEATURE:

NAME/KEY: misc feature

LOCATION: 893..898

OTHER INFORMATION: /note= "Deleted PstI restriction

OTHER INFORMATION: site."

FEATURE:

NAME/KEY: misc feature

LOCATION: 1301..1309

OTHER INFORMATION: /note= "Deleted AlwNI restriction

OTHER INFORMATION: site."

FEATURE:

NAME/KEY: misc feature

LOCATION: 1394..1402

OTHER INFORMATION: /note= "Deleted AlwNI restriction

OTHER INFORMATION: site."

FEATURE:

NAME/KEY: misc signal

LOCATION: 374

OTHER INFORMATION: /note= "RNA start site."

FEATURE:

NAME/KEY: mutation

LOCATION: replace(1671, "")

OTHER INFORMATION: /note= "C to T mutation removes

OTHER INFORMATION: AlwNI site."

FEATURE:

NAME/KEY: CDS

LOCATION: 376..1650

FEATURE:

NAME/KEY: misc feature

LOCATION: 376..414

OTHER INFORMATION: /note= "Bacteriorhodopsin

OTHER INFORMATION: pre-sequence."

FEATURE:

NAME/KEY: terminator

LOCATION: 1672..1674

OTHER INFORMATION: /note= "Bacteriorhodopsin stop

OTHER INFORMATION: codon."

FEATURE:

NAME/KEY: terminator

LOCATION: 1651..1653

OTHER INFORMATION: /note= "Thrombin stop codon."

US-08-767-993-12

Alignment Scores:

Pred. No.: 9.55e-44 Length: 1764

Score: 511.50 Matches: 122

Percent Similarity: 52.07% Conservative: 79

Best Local Similarity: 31.61% Mismatches: 148

Query Match: 23.95% Indels: 37

DB: 3 Gaps: 14

US-09-208-629f-3 (1-407) x US-08-767-993-12 (1-1764)

QY 19 IleLeuIleuValAlaAlaGlyLeuLeuPheLeuProValThr-----ValCys 35

Db 355 ATCTCTGTTAGGTACTGTTGTCATGTTGAGTTATTGCCAACACAGCAGTGGAGGGGTATCG 414

QY 36 GlnSerGlyIleAsnVal-----SerAspAsnSerAlaLysProThrLeuThr 51

Db 415 CAGGCCAGATCCAGCGCTGAGTACAGAGCAGTATGATGACGTCGACGCCACCTTAGAT 474

QY 52 IleLysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAsp 70

Db 475 CCCCGGTCACTTCTCTCAGGAACCCCAATGATAATATGAACCATTTTGGGAGGATGAG 534

QY 71 IleGluGlyTrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysPro 86

Db 535 GAGAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGCAGTCCT 594

QY 87 GluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrlleuArgSerSer 106

Db 595 CTTCAAAAAACACTTCTCGATTCATCTCAGAGTGCCTCCGATATTGACGAGCTCC 654

QY 107 LeuSerThrGlnValIleProAlaIleTyrlleLeuLeuPheValValGlyValProSer 126

Db 655 TGGCTGACACTTTTGTGCCATCTGTGTACACGGAGTGTGTGTAGTCAGCCTCCCACTA 714
Qy 127 AsnIle-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuVal 144
Db 715 AACATCATGCCATCGTTGTGTTCATCTCTGAATAAGAGTCAAGAAG---CCGGCGGTG 771
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 772 GTGTACATGCTGCACCTCGGCAGCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTAA 831
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 832 ATACGCTATTACTTTCCGCGAGTATTCGAGTTTGGGTCTGAATTTGTCTCGTTCGTC 891
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIle 204
Db 892 ACTGCAGCATTTTACTGTAACTATGACGCTCTATCTTGTCTCATGACAGTCATAAGCAT 951
Qy 205 AsnArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLys 221
Db 952 GACCGGTTTCTGGCTGTGTGTATCCCATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1011
Qy 222 ArgSerPheSerLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeu 241
Db 1012 GCTTCTCTCTCT-----TGCTCGCATCTGGGCTTTGGCCATCGCAGGGGTAGTG 1062
Qy 242 ProPheValIleLeuLysGlnTyrHisLeuValHisSerGluIleThrThrCysHis 261
Db 1063 CCTCTGCTCTCAAGAGCAAAACCATCCAGGTGCGCGGCTCAACATCACTACTCTGTCAT 1122
Qy 262 AspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeu 281
Db 1123 GATGTGCTCAATGAAACCTGCTCGAAGGCTACTATGCTACTACTTCTCAGCGCTTCTCT 1182
Qy 282 AlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIle 301
Db 1183 GCTGCTCTCT-----TTTTTGTGCGCTGATCATTTCCACGGTCTGTATGTCTATCAT 1239
Qy 302 HisLysLeuLysSer-----LysAspArgIleTyrPheLeuGly 313
Db 1240 CGATGCTTAGCTTCTCGGAGTTGCCACCGCAGCAAGAAGTCCCGGCTTTG----- 1293
Qy 314 TyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsn 333
Db 1294 TTCTGTGCTGCT-----GTTTCTGCTATCTTCTATCTTCTGCTCGGACCCACAAAC 1347
Qy 334 IleIleLeuValIleHisHisAlaAsnTyrTyrHis-----AsnThrAspSerLeu 351
Db 1348 GTCTCTGATTGGCAATTAC---TCATTCTTCTCACACTTCCACACAGAGGTGCC 1404
Qy 352 TyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProphe 371
Db 1405 TACTTTGCCATCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1464
Qy 372 LeuTyrPheValMetSer 377
Db 1465 ATTACTATTACGCTTCC 1482

RESULT 29
US-07-657-769B-68
; Sequence 68, Application US/07657769B
; Patent No. 5256766
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; TITLE OF INVENTION: RECOMBINANT RECEPTOR AND RELATED
; TITLE OF INVENTION: PHARMACEUTICALS
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IRELL & MANELLA
; STREET: 545 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA

ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/657,769B
FILING DATE: 19910219
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2000-0502.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-327-7250
TELEFAX: 415-327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 3480 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 225..1503
US-07-657-769B-68

Alignment Scores:
Pred. No.: 4,66e-43 Length: 3480
Score: 509.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.83% Indels: 42
DB: 1 Gaps: 14

US-09-208-629F-3 (1-407) x US-07-657-769B-68 (1-3480)

Qy 20 LeuIleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
Db 240 CTGCTCTGTGTGCGCCCTGCTTC-----AGTCTGTGCGGCCCGCTGTG 284
Qy 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
Db 285 TCTGCCCGCACCGCGCCCGCAGCAATCAAAAGCAACAATGCCACTTAGATCCC 344
Qy 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
Db 345 CGGTCACTTTCTCTCAGGAACCCCAATGATAATATGAACCAATTTGGGAGGATGAGGAG 404
Qy 72 GluGlyTyrThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87
Db 405 AAAAATGAAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAAGACGTCCTCT 464
Qy 88 AspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
Db 465 CAAAACAACCTCTCGCATTCATCTCAGAAAGATGCTCCGATATTTGACCACTCTCTGG 524
Qy 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
Db 525 CTGACACTCTTTGTCCCATCTGTGTACACGGAGTGTGTGTGTGTGTGTGTGTGTGTGT 584
Qy 128 Ile-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
Db 585 ATCATGCCATCTGTGTGTCTCATCTCGAAATGAAGTCAAGG---CCGGCGGTGGTG 641
Qy 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
Db 642 TACATGCTGACCTGCGCCACGGCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 701
Qy 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 185


```
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 882 TCCTTCACT-----TGTGGCCATCTGGCTTTGGCCATCGCAGGGGTAGTGCCT 932
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
Db 933 CTGCTCTCAAGAGCAACCATCCAGGTGCCGGGCTCAACATCACTACTGTGATGAT 992
QY 263 ValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeuAla 282
Db 993 GTGCTCAATGAACCCCTGCTCGAAGGTACTATGCTTACTTCTCAGCCTTCTCTGCT 1052
QY 283 PhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHis 302
Db 1053 GTCTTC---TTTTTGTGCGGCTGATCATTTCCACGGCTGTGTTATGTGTCATCATTCGA 1109
QY 303 LysLeuLysSer-----LysAspArgIleTyrPheLeuGlyTyr 314
Db 1110 TGTCTTAGCTCTCCGGAGTGTGCCACCGCAGCAAGAGTCCCGGGCTTTG-----TTC 1163
QY 315 IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 1164 CTGTCAGCTCT-----GTTTTCTGCATCTTCATCATTTGCTTCGGACCCACAAAGTC 1217
QY 335 IleLeuValIleHisHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr 352
Db 1218 CTCTGATTGCGGATTAC---TCATTCCTTTTCACACTTCCACACAGAGGCTGCCTAC 1274
QY 353 PheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu 372
Db 1275 TTTGCTTACTCTCTGCTGTGTGTGTCAGCAGCATAGCTCGTGCTACGCCCCCTAAAT 1334
QY 373 TyrPheValMetSer 377
Db 1335 TACTATTACGCTTCC 1349
```

RESULT 31

```
US-08-475-263-219
; Sequence 219, Application US/08475263
; Patent No. 5759994
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/475,263
; APPLICATION NUMBER: US/08/475,263
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-08-475-263-219

Alignment Scores:
Pred. No.: 4,66e-43 Length: 3480
Score: 509.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.83% Indels: 42
DB: 1 Gaps: 14

US-09-208-629f-3 (1-407) x US-08-475-263-219 (1-3480)

QY 20 LeuIleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
Db 240 CTGCTGCTGCTGGCCGCTCTTC-----AGTCTGTGCGCCCGCTGTTG 284
QY 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
Db 285 TCTGCCCGCACCCGGCCGCGCAGCCAGAAATCAAAAGCAACAATGCCACCTTAGATCCC 344
QY 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
Db 345 CGGTCTATTTCTCTCAGGAACCCCAATGATATGAACCATTTTGGGAGGATGAGGAG 404
QY 72 GluGlyTyrThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87
Db 405 AAAAATGAAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAGCAGCTCCTCT 464
QY 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
Db 465 CAAAAACAATCTCTCGATTCATCTCAGAAAGATGCTCCGGATATTGTGACCACTCTCG 524
QY 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
Db 525 CTCACACTCTTTGTCCTCATCTGTGTACACCGGAGTGTGTGTAGTCAGCTCCCAATAAC 584
QY 128 Ile-----ValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
Db 585 ATCATGGCCATCTGTGTTCATCTCTGAAATGAAGGTCAAGAAG---CGGCGGTGGTG 641
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
Db 642 TACATGCTCACCTCGCCACGCGCAGATGTGCTGTTGTGTCTGTCTCTCCCTTTAAGATC 701
QY 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThr 185
Db 702 AGCTATTACTTTTCGGCAGTGAATGGCAGTTTGGGTCTGAATTGTGCTGCTGCTCACT 761
QY 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsn 205
Db 762 GCAGCATTTTACTGTAAATGATGACGCTCTATCTCTCATGACAGTCATAACCATGAC 821
QY 206 ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg 222
Db 822 CGGTTTCTGGCTGTGCTGATATCCCATGCGAGTCCCTCTCTCTGGGCTACTCTGGAAGG 881
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 882 TCCTTCACT-----TGTCTGGCCATCTGGGCTTTGGCCATCCAGGGGTAGTGCCT 932
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
Db 933 CTGCTCTCAAGAGCAACCATCCAGGTGCCGGGCTCAACATCACTACTCTGTCATGAT 992
QY 263 ValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeuAla 282
Db 993 GTGCTCAATGAACCCCTGCTCGAAGGCTACTATGCTCTACTTCTCTCAGCCTTCTCTGCT 1052
```

283 PhePheGlyPheLeuValIlePheCysTyrThrThrLeuIleHis 302
1053 GTCTTC---TITTTTGGCGCGTATCATTTCCAGGCTGTGTATGTGCTATCATTCGA 1109
303 LysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
1110 TGTCTTAGCTCTCCGCGAGTTGCCAACCAGCAAGAGTCCCGGGCTTTG-----TTC 1163
315 IleLysAlaValLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
1164 CTGTCAGCTGCT-----GTTTCTGTCATCTTCATCATTTTGTTCGACCCCAACGTC 1217
335 IleLeuValIleHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr 352
1218 CTCCTGATGGCATTAC---TCATTCCTTCTCAGCTCCACACAGAGGTCCTAC 1274
353 PheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu 372
1275 TTTGCCTACCTCCTCTGTGTCTGTGTCAGCAGCATAGCTCGTGCATCGACCCCTAATT 1334
373 TyrPheValMetSer 377
1335 TACTATTACGCTTCC 1349

RESULT 32
US-08-485-886-219
; Sequence 219, Application US/08485886
; Patent No. 5798248
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,886
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,184
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-08-485-886-219

Alignment Scores:

Pred. No.: 4,66e-43 Length: 3480
Score: 509.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.83% Indels: 42
DB: 1 Gaps: 14

US-09-208-629f-3 (1-407) x US-08-485-886-219 (1-3480)

QY 20 LeuIleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
DB 240 CTGCTGCTGCTGCGCGCTGCTTC-----AGTCTGTGCGCGCGCTGTGTTG 284
QY 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
DB 285 TGTGCCCGCACCGCGCGCGCGCGAGTCAAAAGCAACAAATGCCACCTTAGATCCC 344
QY 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
DB 345 CGGTCACTTCTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAGGATGAGGAG 404
QY 72 GluGlyTyrThrGlyAlaThr-----ThrIleLysAlaGluCysProGlu 87
DB 405 AAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGCATCTCTT 464
QY 88 AspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
DB 465 CAAAACAACTCTCGCATTCATCTCAGAGATGCCCTCGGATATTTTGACCGACTCCTGG 524
QY 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
DB 525 CTCACACTCTTGTGCCCATCTGTGTACACCGGAGTGTGTGTAGTCAGCTCCCACTAAC 584
QY 128 Ile-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
DB 585 ATCATGCCCATGTTGTGTTCATCTCTGAAATGAAGTCAAGAAG---CCGGCGGTGGTG 641
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
DB 642 TACATGCTGCACCTGGCCAGCGAGATGTGTGTGTCTGTCTCTCCCTTTAGATC 701
QY 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThr 185
DB 702 AGCTATTACTTTTCCGCGCAGTGTGGCAGTTTGGGTCTGAATGTGTGCTCGTCACT 761
QY 186 valValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsn 205
DB 762 GCAGCATTTTACTGTAAACATGTACGCTCTATCTCTCATGACAGTCAATGACATGAC 821
QY 206 ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg 222
DB 822 CGGTTTCTGGCTGTGTGTATCCATGCGAGTCCCTCTCTGGGTACTCTGGGAAGGCT 881
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
DB 882 TCCTTCAC-----TGTCTGGCCATCTGGCTTTGGCCCATCCAGGGGTAGTCCCT 932
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAsp 282
DB 933 CTGCTCCTCAAGGAGCAACCATCCAGGTGCCGGCTCAACATCATCACTACCTCATGAT 992
QY 263 valValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAla 282
DB 993 GTGCTCAATGAACCCCTGTGCGAAGGCTACTATGCGCTACTACTTCTCAGCCCTCTCTGCT 1052
QY 283 PhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeuIleHis 302
DB 1053 GTCTTC---TITTTTGGCGCGTATCATTTCCAGGCTGTGTATGTGTCTATCATTCGA 1109
QY 303 LysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
DB 1110 TGTCTTAGCTCTCCGCGAGTTGCCAACCAGCAAGAGTCCCGGGCTTTG-----TTC 1163

Qy 315 IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 1164 CTGTCAGCTGCT-----GTTTCTGCATCTTCATCATTTGCTTGGACCCACAAAGCTC 1217
 Qy 335 IleLeuValIleHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr 352
 Db 1218 CTCCTGATCGGCATTAC---TCATTCTTCTCAGACATAGCTCGTCATCGACCCCTTAATT 1274
 Qy 353 PheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu 372
 Db 1275 TTTCCTACTCTCTCTGCTGCTGTCAGCAGCATAGCTCGTCATCGACCCCTTAATT 1334
 Qy 373 TyrPheValMetSer 377
 Db 1335 TACTATTACGCTTCC 1349

RESULT 33

US-08-477-362-219
 ; Sequence 219, Application US/08477362
 ; Patent No. 5849507
 ; GENERAL INFORMATION:
 ; APPLICANT: COUGHLIN, SHAUN R.
 ; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
 ; TITLE OF INVENTION: RELATED PHARMACEUTICALS
 ; NUMBER OF SEQUENCES: 223
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/477,362
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/789,184
 ; FILING DATE: 07-NOV-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 22000-20502.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 813-5600
 ; TELEFAX: (415) 494-0792
 ; TELEX: 34-0154
 ; INFORMATION FOR SEQ ID NO: 219:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3480 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 225..1499
 ; US-08-477-362-219

Alignment Scores:
 Pred. No.: 4,66e-43
 Score: 509.00
 Percent Similarity: 52.73%
 Best Local Similarity: 31.17%
 Query Match: 23.83%
 DB: 2

Length: 3480
 Matches: 120
 Conservative: 83
 Mismatches: 140
 Indels: 42
 Gaps: 14

US-09-208-629F-3 (1-407) x US-08-477-362-219 (1-3480)
 Qy 20 LeuIleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
 Db 240 CTGCTCTGGTGGCGCCTGCTTC-----AGTCTGTGGCGCGCGCTGTTG 284
 Qy 40 AsnValSerAspAsnSerAlaIlysPro-----ThrLeuThrIle 52
 Db 285 TCTGCCCGCACCGCGCGCGCGAGCAATCAAAAGCAACAATGCCACCTTAGATCCC 344
 Qy 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluCluPheProLeuSerAspIle 71
 Db 345 CGGTCAATTTCTCTCAGAACCCCAATGATAATGAACCAATTTTGGGAGGATAGGAG 404
 Qy 72 GluGlyTyrThrGlyAlaThr-----ThrIleLysAlaGluCysProGlu 87
 Db 405 AAAAATGAAAGTGGGTAACTGAATACAGATAGTCTCCATCAATAAAACGACGCTCTT 464
 Qy 88 AspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
 Db 465 CAAAAACAACCTCTCTCATTTCTCAGAGATGCTCCGGATATTTGACACGCTCTCTGG 524
 Qy 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
 Db 525 CTGACACTCTTTGTCCTCCATCTGTGTACACCGGAGTGTGTGTAGTACGCTCCCACTAAAC 584
 Qy 128 Ile-----ValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
 Db 585 ATCATGGCCATCGTTGTGTTTCCTGAAATGAAGTCAAGAG---CCGGCGGTGGTG 641
 Qy 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
 Db 642 TACATGCTGCACCTGGCCACGCGAGATGTGTTGTGTCTGTCTGCTCCCTTTAAGATC 701
 Qy 166 AlaTyrHisLeuAsnGlyAsnAsnThrValPheGlyGluValMetCysArgIleThrThr 185
 Db 702 AGCTATTACTTTTCGGCAGTGTGGCTGCAATTTGTGCTGCTGCTGCTGCTGCTGCTG 761
 Qy 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsn 205
 Db 762 GCAGCATTTTACTGTAAACATGACGCTCTAICTGCTCATGACAGTCAATGACATGAC 821
 Qy 206 ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg 222
 Db 822 CGGTTTCTGGCTGTGTGTATCCATGACAGTCCCTCTCTCTGGCGTACTCTGGAAGGCT 881
 Qy 223 SerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeuPro 242
 Db 882 TCTTTCACCT-----TGTCTGGCCATCTGGGCTTTGGCCATCGCAGGGGTAGTGCCT 932
 Qy 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAsp 262
 Db 933 CTGCTCTCAAGGAGCAACCATCCAGGTGCCGGGCTCAACATCACTACTCTGTCATGAT 992
 Qy 263 ValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAla 282
 Db 993 GTGCTCAATGAAACCTCTCGAAGGCTACTATGCTCTACTTCTCTCAGCCTTCTCTGCT 1052
 Qy 283 PhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHis 302
 Db 1053 GTCTTC---TTTTTGTGGCGGTGATCATTTCCACGGTCTGTATGCTGTATCATCATCGA 1109
 Qy 303 LysLeuLysSer-----LysAspArgIleIleTrpLeuGlyTyr 314
 Db 1110 TGTCTTAGCTCTTCCCGAGTTCGCAACCGCAGCAAGAGTCCCGGCTTTG-----TTC 1163
 Qy 315 IleLysAlaValLeuLeuIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 1164 CTGTCAGCTGCT-----GTTTCTGCATCTTCATCATTTGCTTGGACCCACAAAGCTC 1217
 Qy 335 IleLeuValIleHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr 352
 Db 1218 CTCTGATTCGCGATTAC---TCATTCTTCTCAGACATAGCTCGTCATCGACCCCTTAATT 1274

RESULT 35

US-08-473-489A-219
; Sequence 219, Application US/08473489A
; Patent No. 6024936
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,489A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/789,184
; FILING DATE: 1991-11-07
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-08-473-489A-219

Alignment Scores:
Pred. No.: 4,66e-43 Length: 3480
Score: 509.00 Matches: 120
Percent Similarity: 52.73% Conservative: 83
Best Local Similarity: 31.17% Mismatches: 140
Query Match: 23.83% Indels: 42
DB: 3 Gaps: 14

US-09-208-629f-3 (1-407) x US-08-473-489A-219 (1-3480)

QY 20 LeuileuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39
Db 240 CTGCTGCTGGTGGCCGCTGCTTC-----AGTCTGCGCGCGCGCTGTG 284
QY 40 AsnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52
Db 285 TCTGCCCGCCCGCGCGCGCGCGAGATCAAAAGCAACAATGCCACCTTAGATCCC 344
QY 53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
Db 345 CGGTCAATTTCTCTCAGGAACCCCAATGATAAATATCAACATTTTGGGAGGATGAGGAG 404
QY 72 GluGlyTrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87
Db 72 GluGlyTrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 87

Db 405 AAAAATGAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAAGCAGCTCTCTT 464
QY 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
Db 465 CAAAACAACTTCTCCTGCATTCATCTCAGAAGATGCGCTCCGGATATTTTGACCAGCTCTCGG 524
QY 108 SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn 127
Db 525 CTGACACTCTTTTGTCCCATCTGTGTACACCGGAGTGTGTGTAGTCAGCTCCCAACAAAC 584
QY 128 Ile-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIle 145
Db 585 ATCATGGCCATCTGTGTTCATCTCGAAATCAAGGTCAAGAG---CGGCGGGTGGT 641
QY 146 PheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIle 165
Db 642 TACATGCTGCACCTGGCCACGCGAGATGTCTGTGTCTGTCTGCTCCCTTTAAGATC 701
QY 166 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThr 185
Db 702 AGCTATTACTTTTCCGCGCAGTGTGTCAGTGTGGGTCTGAATTGTGTCTGCTGCTCACT 761
QY 186 ValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsn 205
Db 762 GCAGCAATTTTACGTAAACATGACGCCCTCTATCTTCATGACAGTCATAAGCAATGAC 821
QY 206 ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLysArg 222
Db 822 CGGTTTCTGGCTGTGTGTATCCCATGCAGTCCCTCTCTCGGGGTACTCTGGGAAGGGCT 881
QY 223 SerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPro 242
Db 882 TCCTTCACT-----TGCTGGCCATCTGGCTTTGGCCATCGCAGGGGTAGTGCCT 932
QY 243 PheValIleLeuLysGlnGluTyrHisLeuValHisSerCluileThrCysHisAsp 262
Db 933 CTCGCTCTCAAGGAGCAACCATCCAGGTCCCGGGCTCAACATCACTACCTGTCTCAT 992
QY 263 ValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeuAla 282
Db 993 GTGCTCAATGAACCTGTGCGAAGGTACTATGCTACTTCTCAGCCCTTCTCTGCT 1052
QY 283 PhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHis 302
Db 1053 GTCTTC---TTTTTGTGCGCTGATCAATTCACGGTCTGTATGTGTCTATCATTCGA 1109
QY 303 LysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 1110 TGTCTTAGCTCTTCCGAGTTCGCAACCGCAGCAAGAGTCCGCGGCTTTG-----TTC 1163
QY 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 1164 CTGTGAGCTGCT-----GTTTCTGCATCTTCATCATTTGTCGCGACCAACAGCTC 1217
QY 335 IleLeuValIleHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr 352
Db 1218 CTCCTGATTGGCATTAC---TCATTCCTTCTCACACTTCCACCAACAGAGGCTGCCTAC 1274
QY 353 PheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPropheLeu 372
Db 1275 TTGCTACTCTCTGT 1334
QY 373 TyrPheValMetSer 377
Db 1335 TACTATTACGCTTCC 1349

RESULT 36

US-08-485-695-219
; Sequence 219, Application US/08485695
; Patent No. 6124101
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.

[illegible]

Search completed: June 29, 2003, 10:16:58
Job time : 102.944 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model.

Run on: June 29, 2003, 07:26:54 ; Search time 192.264 Seconds
(without alignments)
4767.217 Million cell updates/sec

Title: US-09-208-629F-3

Perfect score: 2136

Sequence: 1 TLTYXHPVAGSQDKMKIL.....AMARPLXRRDRDIWEDIHAW 407

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09208629/runat_24062003_101625_12912/app_query.fasta_1.1166
-DB-N Geneseq 101002 -QWPT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09208629 @CGN 1 1 275 @runat_24062003_101625_12912 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 101002: *
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2121	99.3	1224	19 AAV07372	Mouse protease-act
2	2121	99.3	2409	21 AA250774	Human protease act
3	1443.5	67.6	1224	19 AAV07374	Human protease-act
4	1443.5	67.6	1830	21 AAF21430	Human low adenosin
5	1443.5	67.6	1830	21 AAA35308	Human adenosine re
6	1443.5	67.6	6546	21 AAF21433	Human low adenosin
7	1443.5	67.6	6546	21 AAA35311	Human adenosine re
8	13069	64.1	1102	19 AAV07375	Human protease-act
9	13005	60.9	1124	19 AAV07373	Mouse protease-act
10	584	27.3	2732	16 AAQ84559	Murine C140 recept
11	584	27.3	2732	17 AAQ84557	Murine C140 recept
12	574.5	26.9	1477	16 AAQ84557	Murine C140 recept
13	572.5	26.8	1477	17 AAQ84558	Murine C140 recept
14	563	26.4	1255	16 AAQ84558	Human C140 recepto
15	563	26.4	1255	17 AAQ84558	Human C140 recepto
16	563	26.4	1289	21 AA250773	Human protease act
17	563	26.4	1451	21 AAC60319	Human PAR-2 DNA
18	551	25.8	1414	16 AAQ84560	Human C140 recept
19	551	25.8	1414	17 AAT32039	Human C140 recept
20	514.5	24.1	1312	18 AAT93367	Mouse thrombin rec
21	511.5	23.9	1764	15 AAQ73590	Fragment of the hu
22	509	23.8	3472	20 AA232191	Human thrombin rec
23	509	23.8	3480	21 AA250771	Human thrombin rec
24	507.5	23.8	1534	20 AAQ90981	Human protease-act
25	507.5	23.8	1534	21 AAF21431	Human low adenosin
26	507.5	23.8	1534	21 AAA35309	Human adenosine re
27	507.5	23.8	4895	20 AA232747	Human protease-act
28	507.5	23.8	4895	21 AA250775	Human protease act
29	507.5	23.8	4925	22 AAH43632	Human PAR4 coding
30	502	23.5	2910	18 AAT62461	Thr-GPAL fusion ge
31	502	23.5	3182	21 AAF21432	Human low adenosin
32	502	23.5	3182	21 AAA35310	Human adenosine re
33	501	23.5	1278	24 ABR70889	Human cDNA encodin
34	500	23.4	2137	22 AAH51006	Human nGPCR5 codin
35	497	23.3	1955	21 AA299588	cDNA encoding a hu
36	494	23.1	1080	21 AAD01131	Human orphan G-pro
37	494	23.1	1080	21 AAA46032	Human G-protein co
38	494	23.1	1361	20 AAQ90979	Mouse protease-act
39	493	23.1	1080	22 AAF57649	Human G-protein co
40	493	23.1	1080	23 AAI64231	G-protein coupled
41	493	23.1	1192	21 AAA14828	cDNA encoding a hu
42	491	23.0	3480	13 AAQ28568	Human thrombin rec
43	485.5	22.7	1209	24 ABR70887	Human cDNA encodin
44	484.5	22.7	1116	24 ABR70888	Human cDNA encodin
45	471.5	22.1	2588	19 AAV40372	Mouse G-protein co

ALIGNMENTS

RESULT 1
AAV07372
ID AAV07372 standard; cDNA; 1224 BP.
XX
AC AAV07372;
XX
DT 12-OCT-1998 (first entry)
XX
DE Mouse protease-activated receptor 3 (PAR3) cDNA.
XX
KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
KW G-protein coupled receptor; agonist; antagonist; thrombosin;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis; ds.
XX
OS Mus sp.
XX

FH Key Location/Qualifiers
 FT CDS 51..1160
 FT /*tag= a
 XX WO9818456-A1.
 FN 07-MAY-1998.
 PD 29-OCT-1997; 97WO-US19732.
 PF 30-OCT-1996; 96US-0742440.
 PR (REGC) UNIV CALIFORNIA.
 XX Connolly A, Coughlin SR, Ishihara H;
 XX WPI; 1998-271905/24.
 DR P-PSDB; AAW51405.
 XX
 PT DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.
 XX
 PS Claim 3; Page 38-39; 74pp; English.
 XX
 CC This cDNA clone codes for mouse protease-activated receptor 3
 CC (PAR3) (see AAW51405), a cell surface protein which is specifically
 CC activated by thrombin or a thrombin agonist, thereby activating
 CC signalling events such as phosphoinositide hydrolysis, calcium ion
 CC efflux and platelet aggregation. A cDNA library was established
 CC from rat platelets and amplified (primers given in AAV07376-77). A
 CC product was identified that was expected to encode a new G-protein
 CC coupled receptor, related to PAR1 or 2. This was used to isolate
 CC human and murine DNA by a combination of PCR and hybridisation
 CC techniques. Murine cDNA, murine genomic DNA (see AAV07373), human
 CC cDNA (see AAV07374) and human genomic DNA (see AAV07375) sequences are
 CC provided, as well as the murine and human (see AAW51406) PAR3
 CC proteins. Also claimed are vectors, host cells and an assay
 CC device. Host cells are used to screen compounds for their ability
 CC to act as agonists or antagonists of the effects of thrombin-PAR3
 CC interaction. Agonists are used to treat wounds, thrombosis,
 CC atherosclerosis, restenosis, inflammation and other thrombin
 CC activated disorders. Antagonists (see AAW51415-21) are used to
 CC control blood coagulation and thereby to treat heart attack and
 CC stroke. They also mediate inflammatory and proliferative responses
 CC to injury as occur in wound healing, atherosclerosis, restenosis,
 CC pulmonary inflammation (ARDS) and glomerulosclerosis.
 XX
 SQ Sequence 1224 BP; 299 A; 341 C; 239 G; 345 T; 0 other;
 Alignment Scores:
 Pred. No.: 3,15e-188 Length: 1224
 Score: 2121.00 Matches: 404
 Percent Similarity: 99.26% Conservative: 0
 Best Local Similarity: 99.26% Mismatches: 3
 Query Match: 99.30% Indels: 0
 DB: 19 Gaps: 0
 US-09-208-629F-3 (1-407) x AAV07372 (1-1224)
 Qy 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
 Db 3 ACTTTGTATACTTAACAACATCCTGTAGCGGGTCTCAGGACATCAAGATGAAATCCTT 62
 Qy 21 IleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsn 40
 Db 63 ATCTTGGTGTGAGCTGGGCTCTGTTCTGCGAGTCACTGTTGCCAAAGTGGCATAAAT 122
 Qy 41 ValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGln 60
 Db 123 GTTTCAGACAACTCAGCAACCCACCTTAATTAAGAGTTTAAATGGGGGTCGCCAA 182
 Qy 61 AsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThr 80

Db 183 AATACCTTTGAAGAAATTCCTCACTTTCTGACATAGAGGCTGGACAGAGCCACCAACT 242
 Qy 81 IleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIle 100
 Db 243 ATAAAGCGGAGTGTCCCGAGGACAGTATTTCAACTCTCCACGTGAATATATGTACCAT 302
 Qy 101 GLYTrLeuArgSerSerLeuSerThrGlnValIleProAlaIleIleValIleLeuLeuPhe 120
 Db 303 GGATACCTGAGAAGTTCCTTAACTACCAAGTGATACCTGCCATCTATATCTCTGCTTT 362
 Qy 121 ValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSer 140
 Db 363 GTGGTTGGTGTACCATCCAAACATCGTACCTGTGAAACTCTCTCTTAAGAGCAAAATCC 422
 Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
 Db 423 ATCAGTCTGGTCTATCTTTTACACCAACCTGGGCATCGCAGATCTCTTTTCTGTGCACA 482
 Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnMetTyrCysAlaIleLeuLeuThr 180
 Db 483 CTGCCATTTAAGATCGCTTACCATCTCAATGGCAACAACCTGGGTATTTGGCGAGGTCA 542
 Qy 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
 Db 543 TGCCGGATCACCACGGTCTGTTTCTACGGCAACATGATCTGCGCTATCTCTGATCTCACT 602
 Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
 Db 603 TGCATGGGATCAACCGCTACCTGGCCAGCGCTACCTTTACATACCAAGAGTGGCC 662
 Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
 Db 663 AACGGAGCTTCTCTTGTCTCATGTGTGCATAGTGTGGTCACTGGTTTCTTATACATG 722
 Qy 241 LeuProPheValIleLeuLysGlnLysLeuValHisLeuValHisSerGluLeuThrCys 260
 Db 723 CTGCCCTTTGTCTCTCTGAAGCAGGAGTACCACCTCGTCCACTCAGAGATCACCACCTGC 782
 Qy 261 HisAspValValAspIleCysGluSerProSerSerPheArgPheTyrTyrPheValSer 280
 Db 783 CACGATGTCTGCACGCGTGGAGTCCCATCATCTCTCGATTTCTACTACTTCGTCTCC 842
 Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
 Db 843 TTAGCATTTCTTGGGTCTCTCATCCGTTTGTGATCATCATCTCTGTACAGACTCTC 902
 Qy 301 IleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 320
 Db 903 ATCCACAACTTAATCAAAAGGATCGGATATGGCTGGGTACATCAAGGCGCTCTCTCCTC 962
 Qy 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
 Db 963 ATCCTTGTGATTTTCAAAATTTGCTTTGGCCCCCACCACCAATCATACTCGTAATCCACCAT 1022
 Qy 341 AlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCys 360
 Db 1023 GCCAACTACTACTACCACAAATACCGACAGCTGTACTTTATGTATCTTATGTCTGTGC 1082
 Qy 361 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValVal 380
 Db 1083 CTGGGAGCGCTGAATAGTCTGCCTAGATCCATTCCTTTACTTTCTGTCATGTCTGAAGTGTGA 1142
 Qy 381 AspGlnLeuAsnPro***SerAlaMetAlaAaGProLeu***ArgProArgArgAspIle 400
 Db 1143 GATCAGCTTAATCCTTTAGTGGCAATGGCAAGACCACCTTTAGAGACCAAGGAGATATC 1202
 Qy 401 TrpGluAspIleHisAlaTrp 407
 Db 1203 TGGGAAGACATACATGCTTGG 1223
 RESULT 2
 AAZ50774

ID AAZ50774 standard; DNA; 2409 BP.
 AC AAZ50774;
 XX
 DT 31-MAY-2000 (first entry)
 XX
 DE Human protease activated receptor-3 DNA.
 XX
 KW Human; PAR-3; protease activated receptor;
 KW antisense molecule; PAR antibody; cytostatic; therapeutic;
 KW metastatic tumour cell; placental implantation; invasive cell; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200008150-A1.
 XX
 PD 17-FEB-2000.
 XX
 PF 05-FEB-1999; 99WO-1100079.
 XX
 PR 07-AUG-1998; 98IL-0125698.
 XX
 PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
 XX
 PI Bar-Shavit R;
 XX
 DR WPI; 2000-205706/18.
 XX
 PT Treating metastatic tumor cells useful for treating disorders involving
 PT placenta implantation in a female comprises administration of an
 PT antisense molecule complementary to an RNA sequence of a protease
 PT activated receptor protein -
 XX
 PS Example 3; Fig 10; 46pp; English.
 XX
 CC The patent discloses a method to treat metastatic tumour cells using
 CC an antisense molecule comprising a polynucleotide complementary to an
 CC RNA sequence of a protease activated receptor (PAR) protein, or an
 CC antibody capable of binding to a PAR protein. The antisense molecules and
 CC antibodies of PAR protein are also used to treat disorders associated
 CC with implantation of placenta. The present sequence is a human PAR-3
 CC DNA used for producing antisense molecules for treating invasive
 CC cells.
 XX
 SQ Sequence 2409 BP; 598 A; 652 C; 454 G; 703 T; 2 other;
 Alignment Scores:
 Pred. No.: 7.63e-188 Length: 2409
 Score: 2121.00 Matches: 404
 Percent Similarity: 99.26% Conservative: 0
 Best Local Similarity: 99.26% Mismatches: 3
 Query Match: 99.30% Indels: 0
 DB: 21 Gaps: 0
 US-09-208-629F-3 (1-407) x AAZ50774 (1-2409)
 QY 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
 DB 105 ACITTTGTACTTAACAAACATCTCTAGCCGGTCTCAGGACATCAAGATGAATTCCTT 164
 QY 21 IleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsn 40
 DB 165 ATCTTGGTTCAGCTGGGCTGCTTCTCCAGTCACTGTTGCCAAAGTGCATAAAT 224
 QY 41 ValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyLysProGln 60
 DB 225 GTTTCAGCAACATCAGCAAGCCAACTTAACATTAAAGAGTTTAAATGGGGTCCCAA 284
 QY 61 AsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThr 80
 DB 285 AATACCTTTGAGAAATTCCTCCACTTTCTGACATAGAGGGCTGGACAGGAGCCACCAACT 344
 QY 81 IleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIle 100

Db 345 ATAAAAGCGAGGTCCCGAGGACAGTATTTCACCTCTCCAGTGAATAATGTACCATA 404
 QY 101 GlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPhe 120
 Db 405 GGATACCTGAGAAGTTCTTAACTACCAAGTATACCTGCCATCTATATCTCTGCTGTTT 464
 QY 121 ValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSer 140
 Db 465 GTGGTTGGTGTACCATCCAAACATCGTGACCTGTGGAAACTCTCTTAAGGACCAATCC 524
 QY 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
 Db 525 ATCAGTCTGTCTATCTTTTACACCAACCTGGCCATCGCAGATCTCCTTTTCTGTGCACA 584
 QY 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
 Db 585 CTGCCATTTAAGATCGCTTACCATCTCAATGGCAACAACTGGGTATTTGGCGAGGTCATG 644
 QY 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
 Db 645 TGCCGGATCACCAACGGTCTGTTTCTACGGCAACATGCTACTGGCTATCTCTGATCTCACT 704
 QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
 Db 705 TGCATGGCATCAACCGCTACCTGGCCACGGCTCACCTTTCACATACCAAGAGCTGCC 764
 QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
 Db 765 AAACGAGCTTCTCTTGTCTCATGTGTGCATAGTGTGGTCAATGGTTTCTTATACATG 824
 QY 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
 Db 825 CTGCCCTTTGTCTATCTCTGAAGCAGGAGTACCACCTCTCTCCACTCAGAGATCACCACTGC 884
 QY 261 HisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSer 280
 Db 885 CACGATGTCTGCACGCGTGCAGAGTCCCATCATCTCTCCGATCTCTACTACTTCTGCTCC 944
 QY 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
 Db 945 TTAGCATCTTTGGGGTCTCTCATCCGTTTGTGATCATCATCTTCTGTACAGACTCTC 1004
 QY 301 IleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 320
 Db 1005 ATCCCAAACTTAATCAAAAGGATCGGATATGCTGGGTACATCAAGCCGCTCTCTCCTC 1064
 QY 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
 Db 1065 ATCCTTGTGATTTTCACAATTTGCTTTGGCCCCCACCACCAACATCATACTCGTAATCCACCAT 1124
 QY 341 AlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleLeuAlaLeuCys 360
 Db 1125 GCCAACTACTACTACCACAAATACCGACGCTGTACTTTATGTATCTTATGTCTGTGTC 1184
 QY 361 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValVal 380
 Db 1185 CTGGGGAGCGCTGATAGCTGCCTAGATCCATCTCTTACTTGTCTATGTCGAAGTGTGA 1244
 QY 381 AspGlnLeuAsnPro***SerAlaMetAlaArgProLeu***ArgProArgArgAspIle 400
 Db 1245 GATCAGCTTAATCTCTTAGTCGGCAATGGCAAGACCACTTTAGAGACCAAGGAGATATC 1304
 QY 401 TrpGluAspIleHisAlaTrp 407
 Db 1305 TGGGAAGACATACATGCTTGG 1325
 RESULT 3
 AAV07374
 ID AAV07374 standard; cDNA; 1224 BP.
 XX
 AC AAV07374;
 XX

DT 12-OCT-1998 (first entry)
 XX Human protease-activated receptor 3 (PAR3) cDNA.
 DE
 XX
 XX Protease-activated receptor 3; PAR3; thrombin receptor; human;
 KW G-protein coupled receptor; agonist; antagonist; thrombosis;
 KW atherosclerosis; restenosis; inflammation; blood coagulation;
 KW blood clotting; heart attack; stroke; wound healing;
 KW adult respiratory distress syndrome; glomerulosclerosis; ds.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 58..1182
 FT CDS /tag= a
 FT
 XX
 PN W09818456-A1.
 XX
 PD 07-MAY-1998.
 XX
 PF 29-OCT-1997; 97WO-US19732.
 XX
 XX 30-OCT-1996; 96US-0742440.
 PR
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Connolly A, Coughlin SR, Ishihara H;
 PI
 XX WPI; 1998-271905/24.
 DR P-PSDB; AAW51406.
 DR
 XX
 XX DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.
 FT
 XX
 PS Claim 3; Page 40-41; 74pp; English.
 XX
 CC This cDNA clone codes for mouse protease-activated receptor 3
 CC (PAR3) (see AAW51406), a cell surface protein which is specifically
 CC activated by thrombin or a thrombin agonist, thereby activating
 CC signalling events such as phosphoinositide hydrolysis, calcium ion
 CC efflux and platelet aggregation. A cDNA library was established
 CC from rat platelets and amplified (primers given in AAV07376-77). A
 CC product was identified that was expected to encode a new G-protein
 CC coupled receptor, related to PAR1 or 2. This was used to isolate
 CC human and murine DNA by a combination of PCR and hybridisation
 CC techniques. Murine cDNA (see AAV07372), murine genomic DNA (see
 CC AAV07373), human cDNA and human genomic DNA (see AAV07375) sequences
 CC are provided, as well as the murine (see AAW51405) and human PAR3
 CC proteins. Also claimed are vectors, host cells and an assay
 CC device. Host cells are used to screen compounds for their ability
 CC to act as agonists or antagonists of the effects of thrombin-PAR3
 CC interaction. Agonists are used to treat wounds, thrombosis,
 CC atherosclerosis, restenosis, inflammation and other thrombin
 CC activated disorders. Antagonists (see AAW51415-21) are used to
 CC control blood coagulation and thereby to treat heart attack and
 CC stroke. They also mediate inflammatory and proliferative responses
 CC to injury as occur in wound healing, atherosclerosis, restenosis,
 CC pulmonary inflammation (ARDS) and glomerulosclerosis.
 XX
 SQ Sequence 1224 BP; 303 A; 316 C; 221 G; 384 T; 0 other;
 Alignment Scores:
 Pred. No.: 4,51e-125 Length: 1224
 Score: 1443.50 Matches: 264
 Percent Similarity: 83.74% Conservative: 45
 Best Local Similarity: 71.54% Mismatches: 59
 Query Match: 67.58% Indels: 1
 DB: 19 Gaps: 1
 US-09-208-629F-3 (1-407) x AAV07374 (1-1224)
 Qy 11 GlySerGlnAspIleLysMetLysIleLeuValAlaGlyLeuLeuPheLeu 30

40 GGGACTCAGGTTCATCAAAATGAAAGCCCTCATCTTTGACAGCTGCTGCGCTCTCTGCTTCG 99
 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlaLysProThr 49
 100 TTGCCCATCTTTTGTGAGAGTGGCAATGATAAACAACATGGAAGCAAGCAACC 159
 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
 160 TTACCATTAAGACCTTTCTGAGAGCTCCCAAAATTCCTTTGAAGAGTTCCCTTTCT 219
 70 AspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
 220 GCCTTGAAGAGCTGGACAGGACGACCATTAAGTAAATTAAGTGCCTGCAAGAAAGT 279
 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyLeuArgSerSerLeuSerThr 109
 280 GCTTCACATCTCCATGTGAAATGCTACCATGGGTACCTGACCAAGCTCCCTTAAGTACT 339
 110 GlnValIleProAlaIleTyIleLeuLeuPheValValGlyValProSerAsnIleVal 129
 340 AAATGATACCTGCCATCTACCTCTGGTGTGTAGTTGGTGTCCCGGCCAATGCTGTG 399
 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149
 400 ACCCTGTGATGCTTTCTTCAGGACCATCATCTGTACCATCTGTATTCTACCAAC 459
 150 LeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyHisLeu 169
 460 CTGGCCATTGCAGATTTCTTTTGTGTATACATTTGCCCTTTAAGATAGCTTATCATCTC 519
 170 AsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyr 189
 520 AATGGAAACATGGGTATTTGGAGAGGTCCTGTCCCGGCCACCAAGTATCTCTAT 579
 190 GlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAla 209
 580 GGCAACATGCTACTGCTCCATTTCTGCTCTGCTGCATCAGCATCAACCGTACCTGGCC 639
 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229
 640 ATGCTCCATCTTTCACCTACCGGGCCCTGCCCAAGCACCATCTATGCTTGGTAAATGT 699
 230 GlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlu 249
 700 GGACTGTGTGGCAACAGTTTCTATATATGCTGCCATTTTTCATCTGAAGCAGGAA 759
 250 TyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSer 269
 760 TATTATCTTGTTCAGCCAGACATCACCATCTGCTGCTGCTTTCATCTGAAGCAGTCC 819
 270 ProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIlePro 289
 820 TCATCTCCCTTCCAACTTATTACTTCTCTCTGTCATCTCTGTCATCTTTCGATCTTAAATTC 879
 290 PheValIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerLysAspArg 309
 880 TTGTGCTTATCATCTATGCTATGAGCATCATCCGACACATTAATGCATAGCATCAT 939
 310 IleTrpLeuGlyTyIleLysAlaValLeuLeuValIleLeuValIlePheThrIleCysPhe 329
 940 AGATGTTGTGTATGTTAAGCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 999
 330 AlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAsp 349
 1000 GCTCCAAGCAATATTATTCTTATTATTCACCATGCTAACTACTACTACTACAAACACTGAT 1059
 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
 1060 GGCCTTATATTATATATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1119
 370 ProPheLeuTyrPheValMetSerLys 378

Db 1120 CCATTCCTTTATTTCTCATGTCAAAA 1146
RESULT 4
AAF21430
ID AAF21430 standard; DNA; 1830 BP.
XX AC AAF21430;
XX AC AAF21430;
XX 14-MAR-2001 (first entry)
XX Human low adenosine antisense oligonucleotide related sequence #2997.
XX
XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200062736-A2.
PN
XX
XX 26-OCT-2000.
PD
XX
XX 24-MAR-2000; 2000WO-US08020.
PF
XX
XX 06-APR-1999; 99US-0127958.
PR
XX
XX (UYEC-) UNIV EAST CAROLINA.
PA
XX
XX (NYCE/) NYCE J W.
PA
XX
XX Nyce JW;
PI
XX
XX WPI; 2000-679539/66.
DR
XX
XX Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
XX Disclosure; Page 1414-1415; 1592pp; English.
PS
XX
XX The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and/or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide

CC fragments and antisense oligonucleotides used in the exemplification of
XX the present invention.
SQ Sequence 1830 BP; 473 A; 464 C; 337 G; 556 T; 0 other;
Alignment Scores:
Pred. No.: 7.64e-125 Length: 1830
Score: 1443.50 Matches: 264
Percent Similarity: 83.74% Conservatives: 45
Best Local Similarity: 71.54% Mismatches: 59
Query Match: 67.58% Indels: 1
DB: 21 Gaps: 1
US-09-208-629f-3 (1-407) x AAF21430 (1-1830)
QY 11 GlySerGlnAspIleLysMetLysIleLeuValAlaAlaGlyLeuLeuPheLeu 30
Db 127 GGGACTCAGGTCAATCAAAATGAAAGCCCTCACTTTGGAGCTGCTGGCTCTCTCTCTG 186
QY 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlaLysProThr 49
Db 187 TTGCCCACTTTTGTTCAGAGTGGCATGGAAATGATACAAACAACTTTGGCAAGCCAACC 246
QY 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
Db 247 TTACCCATTAAGACCTTTTCGTGGAGCTCCGCCAAATTCCTTTGAAGAGTTCCCTCTTCT 306
QY 70 AspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
Db 307 GCCTTGGAGGCTGGACAGGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 366
QY 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTrpLeuArgSerSerLeuSerThr 109
Db 367 GCTTTCACATCTCCATGTGAAATGCTACCATGGGTACCTGACGACGACGACGACGACGAC 426
QY 110 GlnValIleProAlaIleTrpIleLeuLeuPheValValGlyValProSerAsnIleVal 129
Db 427 AAACGTATACCTGCCATCTACCTCTCTGTGTTTGTAGTTGGTGTCCCGGCAATGCTGTG 486
QY 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149
Db 487 ACCTGTGTGATGCTTTCTTTCAGGACGACGACGACGACGACGACGACGACGACGACGAC 546
QY 150 LeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTrpHisLeu 169
Db 547 CTGGCCATTCGAGATTTCTTTTGTGTACATTCGCTTAAAGATAGCTTATCATCTC 606
QY 170 AsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyr 189
Db 607 AATGGGAACAACCTGGGTATTTGGAGAGGTCCTGTGCGGGCCACACAGTCATCTCTCTAT 666
QY 190 GlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeuAla 209
Db 667 GGCAACATGTACTGCTCCATTCGCTTTCGCTGTCATCAGCATCAACGCTACCTGGCC 726
QY 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229
Db 727 ATCGTCCATCTTTCACCTACCGGGCTGCCCAACACACACCTATGCTTGGTAACATGT 786
QY 230 GlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlu 249
Db 787 GGACTGCTGGGCAACAGTTTCTTATATATATGCTGCCATTTTTCATCTACTGAAGCAGGAA 846
QY 250 TyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSer 269
Db 847 TATTATCTTGTTCAGCAGACATCACCATCGCATGATGTTCCACACACTTCGGAGTCC 906
QY 270 ProSerSerPheArgPheTyrTrpPheValSerLeuAlaPhePheGlyPheLeuIlePro 289
Db 907 TCATCTCCCTCCAACTCTATTAATCTCTCTCTGCGCATCTCTTGGATCTTCTTAATCCA 966
QY 290 PheValIleIleLeuPheCysTyrThrThrLeuIleHisLeuLysSerLysAspArg 309

Db 967 TTTGTGCTTATCATCTACTGCTATGAGCCATCATCGGACACTTAATGATCATGATCAT 1026
Qy 310 IletirpLeuGlyTyrilleLysAlaValLeuLeuLeuVallePheThrilleCysPhe 329
Db 1027 AGATGGTTGGTATGTTAGGCGAGTCTCTCTATCTCTGATTTTACATTTGCTTT 1086
Qy 330 AlaProThrAenilleLeuValilleHieHialaAsnTyrTyrTyrHisAenThrAsp 349
Db 1087 GCTCCAAGCAATATTATTCTTATTATTACCATGCTTAACCTACTACTACAACAACACTGAT 1146
Qy 350 SerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAenSerCysLeuAsp 369
Db 1147 GCCTTATATTTTATATATCTCATAGCTTTGCGCTGGGTAGTCTTATATAGTCTTAGAT 1206
Qy 370 ProPheLeuTyrPheValMetSerLys 378
Db 1207 CCATCTCTTATTTTCTCAAGTCNAAA 1233
RESULT 5
AAA35308
ID AAA35308 standard; DNA; 1830 BP.
XX
AC AAA35308;
XX
28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:182.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US17712.
XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI; 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -
XX
PS Disclosure; Page 1329; 1343pp; English.

XX
CC The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasize to the lungs, including

CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1680
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.

XX
SQ Sequence 1830 BP; 473 A; 464 C; 337 G; 556 T; 0 other;

Alignment Scores:

Pred. No.: 7,64e-125 Length: 1830
Score: 1443.50 Matches: 264
Percent Similarity: 83.74% Conservative: 45
Best Local Similarity: 71.54% Mismatches: 59
Query Match: 67.58% Indels: 1
DB: 21 Gaps: 1

US-09-208-629F-3 (1-407) x AAA35308 (1-1830)

Qy 11 GlySerGlnAspIleLysMetLysIleLeuValAlaAlaGlyLeuPheLeu 30
Db 127 GGGACTCAGTCATCAAAATGAAAGCCCTCATCTTTGCAGCTGCTGCTCTCTGCTTCG 186
Qy 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlaLysProThr 49
Db 187 TTGCCCACTTTTGTGCAGAGTGCGCATGGAATAATGATACAAACAACTTGGCAAGCCAAAC 246
Qy 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
Db 247 TTACCCATTAAAGACCTTTCTGTGAGCTCCCAAAATTCCTTTTGAAGAGTTCCCTTTTCT 306
Qy 70 AspIleGluGlyTyrThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
Db 307 GCCTTGGAGAGCTGGACAGGAGCCAGGATTAAGTAAATTAAGTGCCTGAAGAAAGT 366
Qy 90 IleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThr 109
Db 367 GCTTCACATCTCCATGTGAAAAATGCTACCACTGGGTACCTGACCAAGCTCTTTAAGTACT 426
Qy 110 GlnValIleProAlaIleTyrIleLeuLeuPheValIleGlyValProSerAsnIleVal 129
Db 427 AAACGTATACCTGCCATCTACCTCTCTGGTGTGTGGTGTCCCGCCCAATGCTGTG 486
Qy 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149
Db 487 ACCCTGTGGATGCTTTTCTTCAGGACCAGATCATCTGTACCATCTATTCTACCAAC 546
Qy 150 LeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 169
Db 547 CTGGCCATTGCGAGATTTCTTTTGTGTATACATTCCTCTTAAAGATAGCTTATCATCTC 606
Qy 170 AsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyr 189
Db 607 AATGGGAACAACCTGGGTATTTGGAGAGGTCTGTGCGGGCCACCAAGTCACTCTCTAT 666
Qy 190 GlyAsnMetTyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAla 209
Db 667 GGCAACATGATGCTCTCCATCTGCTCTTGCCTGCATCAGCATCAACCGCTACCTGGCC 726
Qy 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229
Db 727 ATCGTCCATCTCTTACCTACCGGGGCTGCCCAAGCACACCTATGCTTGGTAAACATGT 786
Qy 230 GlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlu 249
Db 787 GGACTGGTGGGCAACAGATTTCCTTATATATGCTGCCATTTTTCATCTAGAACAGGAA 846

Db 427 AACTGATACCTGCCATCTACCTCTCGTGTGTGTAGTTGGTCCGGCCAACTGCTGTG 486
Qy 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerLeuValIlePheHisThrAsn 149
Db 487 ACCCTGTGGATGCTTTCTTCAGACACAGATCCATCTGTACCATCTGTATCTACACCAAC 546
Qy 150 LeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 169
Db 547 CTGGCCATTCAGATTTCTTTTGTGTGTACATTCCTTTAAGATAGCTTATCATCTC 606
Qy 170 AsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyr 189
Db 607 AATGGGAACAACTGGGTATTTGGAGAGGTCTGTGCGCGGCCACCAAGATCTTCTAT 666
Qy 190 GlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeuAla 209
Db 667 GGCAACATGTACTGCTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726
Qy 210 ThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCys 229
Db 727 ATCGTCCATCTCTTCACCTACCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 786
Qy 230 GlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlu 249
Db 787 GGACTGTGTGGCAACAGTTTCTATATATATGCTGCTGCTGCTGCTGCTGCTGCTG 846
Qy 250 TyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSer 269
Db 847 TATTATCTGTTCAGCCAGACATCACCACCTGCCATGATGTTCAACACACTTGGAGTCC 906
Qy 270 ProSerSerPheArgPheTyrPheValSerLeuAlaPhePheGlyPheLeuIlePro 289
Db 907 TCATCTCCCTTCCAACTCTATTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
Qy 290 PheValIleIlePheCysTyrThrThrLeuIleHisLysLysSerLysAspArg 309
Db 967 TTGTGCTTATCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
Qy 310 IleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIleCysPhe 329
Db 1027 AGATGGTGTGTATGTTAAGCGGAGTCTCTCATCTCTGCTGCTGCTGCTGCTGCTG 1086
Qy 330 AlaProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAsp 349
Db 1087 GCTCCAGCAATATTATTATTATTCACCATGCTAACTACTACTACTACTACTACTACT 1146
Qy 350 SerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 369
Db 1147 GGCTTATATTTATATATCTCATAGCTTTGTGCTGGGTAGTCTTAAATAGTCTTAGAT 1206
Qy 370 ProPheLeuTyrPheValMetSerLys 378
Db 1207 CCATTCCTTATTCTCATGTCAAAA 1233

RESULT 8
AAV07375
ID AAV07375 standard; DNA; 1102 BP.
XX
AC AAV07375;
XX
DT 12-OCT-1998 (first entry)
XX
DE Human protease-activated receptor 3 (PAR3) genomic DNA.
XX
KW Protease-activated receptor 3; PAR3; thrombin receptor; human;
KW G-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis; ds.
OS Homo sapiens.
XX
PN W09818456-A1.

XX 07-MAY-1998.
XX 29-OCT-1997; 97WO-US19732.
XX 30-OCT-1996; 96US-0742440.
XX (REGC) UNIV CALIFORNIA.
XX Connolly A, Coughlin SR, Ishihara H;
XX WPI; 1998-271905/24.
XX
XX DNA encoding protease-activated receptor 3 - for detection of
XX specific agonists and antagonists, potentially useful for treating
XX e.g. thrombosis, atherosclerosis, inflammation etc.
XX
XX Claim 3; Page 41; 74pp; English.
XX
XX This sequence comprises genomic DNA (excluding exon 2) for human
XX protease-activated receptor 3 (PAR3), a cell surface protein which
XX is specifically activated by thrombin or a thrombin agonist,
XX thereby activating signalling events such as phosphoinositide
XX hydrolysis, calcium ion efflux and platelet aggregation. A cDNA
XX library was established from rat platelets and amplified (see
XX AAV07376-77). A product was identified that was expected to encode a
XX new G-protein coupled receptor, related to PAR1 or 2. This was
XX used to isolate human and murine DNA by a combination of PCR and
XX hybridisation techniques. Murine cDNA (see AAV07372), murine
XX genomic DNA (see AAV07373), human cDNA (see AAV07374) and human genomic
XX DNA sequences are provided, as well as the murine (see AAV51405) and
XX human (see AAV51406) PAR3 proteins. Also claimed are vectors, host
XX cells and an assay device. Host cells are used to screen compounds
XX for their ability to act as agonists or antagonists of the effects
XX of thrombin-PAR3 interaction. Agonists are used to treat wounds,
XX thrombosis, atherosclerosis, restenosis, inflammation and other
XX thrombin activated disorders. Antagonists (see AAV51415-21) are used
XX to control blood coagulation and thereby to treat heart attack and
XX stroke. They also mediate inflammatory and proliferative responses
XX to injury as occur in wound healing, atherosclerosis, restenosis,
XX pulmonary inflammation (ARDS) and glomerulosclerosis.
XX
SQ Sequence 1102 BP; 276 A; 286 C; 195 G; 345 T; 0 other;

Alignment Scores:
Pred. No.: 3 46e-118 Length: 1102
Score: 1369.00 Matches: 245
Percent Similarity: 84.96% Conservative: 43
Best Local Similarity: 72.27% Mismatches: 51
Query Match: 64.09% Indels: 0
DB: 19 Gaps: 0

US-09-208-629F-3 (1-407) x AAV07375 (1-1102)

Qy 40 AsnValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyPro 59
Db 13 AATGATACAAACAACTTGGCAAGCCAACTTACCATTAAAGACCTTTCGTGAGCTCCC 72
Qy 60 GlnAsnThrPheGluCluPhePheProLeuSerAspIleGlyTyrThrGlyAlaThrThr 79
Db 73 CCAATTTCTTTTGAAGAGTTCCTCTTCTGCTTGGAGGCTGGACAGGACCCGATT 132
Qy 80 ThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThr 99
Db 133 ACTGTAAAAATTAAAGTGCCTCGAAGAAAGTGTTCATCTCCATGTGAAAAATGTACC 192
Qy 100 IleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeu 119
Db 193 ATGGGGTACTCTGACACAGCTCTTAACTACTAACTGATACCTGCTACCTCTCTGGTG 252
Qy 120 PheValValGlyValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLys 139
Db 253 TTGTAGTTGGTGTCCCGCCCAATGTGTGACCTGTGTGATGCTTTTCTTTCAGGACCA 312

QY	140	SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysVal	159
Db	313	TCCATCTGTACCAGTGTATCTACACCAACCTGGCCATTCGCAGATTTCTTTTGTGTT	372
QY	160	ThrLeuProPheLysIleAlaIlePheHisThrAsnLeuValPheGlyGluVal	179
Db	373	ACATATSCCTTTTAAAGATAGCTTATCATCTCAATGGGAACAACCTGGGTATTGTGAGAGGTC	432
QY	180	MetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeu	199
Db	433	CTGTGCGGGGCCACACAGTCATCTTCTATGGCAACATGCTGCTCCATCTGCTCCCT	492
QY	200	ThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeu	219
Db	493	GCCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCCTTTCACCTACCGGGGCGTG	552
QY	220	ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr	239
Db	553	CCCAAGCACACCTATGCTTGGTAACATGTGGATGGGTGGCAACAGTTTTCTTATAT	612
QY	240	MetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThr	259
Db	613	ATGCTGCCATTTTCATACTGAAGCAGGAATATATCTTGTTCAGCCAGACATCACACC	672
QY	260	CysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheVal	279
Db	673	TGCCATGATGTTCAACAACACTTGCAGATCCTCATCTCCCTTCCAACCTATTACTTCATC	732
QY	280	SerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThr	299
Db	733	TCCTTGGCAATCTTTGGATTCTTAATTCATTTGGCTTATCATCTGCTATGACGCC	792
QY	300	LeuIleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeu	319
Db	793	ATCATCCGGACACTTAATGCATACGATCATAGATGGTTGGTATGTTAAGGGGAGTCTC	852
QY	320	LeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHis	339
Db	853	CTCATCTTGTGATTTTACCATTTCCTTGGCTTTCCTCAAGCAATATTATTCTATTATTAC	912
QY	340	HisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeu	359
Db	913	CATGCTAACTACTACACAACACTGATGGCTTATATTTTATATATCATAGCTTTCG	972
QY	360	CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys	378
Db	973	TGCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTATTTTCTCATGTCAAAA	1029

Qy	29	PheLeuProValThrValCysGlnSerClyIleAsnValSerAspAsnSerAlaLysPro	48
Db	15	TTCTTTTCAATTACA-----GGCATAAATGTTTCAGACAACTCAGCAAAGCCA	62
Qy	49	ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu	68
Db	63	ACCTTAACATTTAAGAGTTTTAATGGGGGTCCCAAAATACCTTTGAAGAATTCNN	118
Qy	69	SerAspIleGluGlyTrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAsp	88
Db	118	-----	111
Qy	89	SerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSer	108
Db	119	NNNNNTACCACTCTCCATGTGAATAATGCTACCATGGGATACCTGAGAAAGTTCTTTAAGT	178
Qy	109	ThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIle	128
Db	179	ACCAAGTGATACCTGCCATCTACATCTCGGTGTTTGATGTGTGACAGCAATC	238
Qy	129	ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThr	148
Db	239	GTGACCCCTGTGGAAACTCTCCTCAAGGACCAAAATCCATCTGCTGGTCATCTTTCCACC	298

PR 30-OCT-1996; 96US-0742440.
XX
XX (REGC) UNIV CALIFORNIA.
PA
XX
XX Connolly A, Coughlin SR, Ishihara H;
PI
XX WPI; 1998-271905/24.
DR
XX
XX DNA encoding protease-activated receptor 3 - for detection of
PT specific agonists and antagonists, potentially useful for treating
PT e.g. thrombosis, atherosclerosis, inflammation etc.
PT
XX
XX Claim 3; Page 39; 74pb; English.
PS

This sequence comprises genomic DNA for mouse protease-activated receptor 3 (PAR3), a cell surface protein which is specifically activated by thrombin or a thrombin agonist, thereby activating signalling events such as phosphoinositide hydrolysis, calcium ion efflux and platelet aggregation. A cDNA library was established from rat platelets and amplified (primers given in AA07376-77). A product was identified that was expected to encode a new G-protein coupled receptor, related to PAR1 or 2. This was used to isolate human and murine DNA by a combination of PCR and hybridisation techniques. Murine cDNA (see AA07372), murine genomic DNA, human cDNA (see AA07374) and human genomic DNA (see AA07375) sequences are provided, as well as the murine (see AA051405) and human (see AA051406) PAR3 proteins. Also claimed are vectors, host cells and an assay device. Host cells are used to screen compounds for their ability to act as agonists or antagonists of the effects of thrombin-PAR3 interaction. Agonists are used to treat wounds, thrombosis, atherosclerosis, restenosis, inflammation and other thrombin activated disorders. Antagonists (see AA051415-21) are used to control blood coagulation and thereby to treat heart attack and stroke. They also mediate inflammatory and proliferative responses to injury as occur in wound healing, atherosclerosis, restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis.

Sequence 1124 BP: 251 A: 314 C: 210 G: 326 T: 23 other: XX

Alignment Scores:		
Pred. No.:	8.64e-112	Length:
Score:	1300.50	Matches:
Percent Similarity:	71.99%	Conservative:
Best Local Similarity:	69.37%	Mismatches:
Query Match:	60.88%	Indels:
DB:	19	Gaps:
		4
		1124

US-09-208-629F-3 (1-407) X AAV07373 (1-1124)

Qy	29	PheLeuProValThrValCysGlnSerClyIleAsnValSerAspAsnSerAlaLysPro	48
Db	15	TTCTTTTCAATTACA-----GGCATAAATGTTTCAGACAACTTCAGCAAGCCA	62
Qy	49	ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu	68
Db	63	ACCTTAACATTTAAGAGTTTTAATGGGGGTCCCAAAATACCTTTGAAGAATTCNN	118
Qy	69	SerAspIleGluGlyTrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAsp	88
Db	118	-----	111
Qy	89	SerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSer	108
Db	119	NNNNNTACCACTCTCCATGTGAATAATGCTACCATGGGATACCTGAGAAAGTTCTTTAAGT	178
Qy	109	ThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIle	128
Db	179	ACCAAGTGATACCTGCCATCTACATCTCGGTGTTTGATGTGTGACAGCAATC	238
Qy	129	ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThr	148
Db	239	GTGACCCCTGTGGAAACTCTCCTCAAGGACCAAAATCCATCTGCTGGTCATCTTTCCACC	298

Qy 149 AsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheIleAlaTyrHi 168
 Db 299 AACCTGGCCATCGCGATCTCTTTCTGTGTACAGCTGCGGTTTAAAGATCNCCTACCA 358
 Qy 168 sLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPh 188
 Db 359 TCTCAATGGCAACAACTGGGTATTTGGGAGGTCTGTGCGGATCACCAGGTCGTTTT 418
 Qy 188 eTyrGlyAsnMetTyrCysAla----IleLeuIleLeuThrCysMetGlyIleAsnArgTy 207
 Db 419 CTACGGCAACATGTACTCGGTANNNTCTGTATCTCACCCTGCATGGCATCAACGCTA 478
 Qy 207 rLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLe 227
 Db 479 CCTGGCCACGGCTACACCTTTTACATACCAAGAGCTGCCAAACGCAGCTTCTCCATGT 538
 Qy 227 uMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIle-LeuL 247
 Db 539 CATGTGTGGCATGGTGTGGTCTATGTTTCTTATACATGCTGCCCTTTGTCTCCNNNA 598
 Qy 247 ySglnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaC 267
 Db 599 AGCAGGAGTACCACCTCGTCCACTCCGAGATCACCACCTGCCAGATGTCTGCGACGGT 658
 Qy 267 ySglnSerProSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheL 287
 Db 659 GCGANTCCCATCATCTCCGATTTCTACTTCTGCTCTCTAGCATTTCTTTGGGTTC 718
 Qy 287 euIleProPheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerL 307
 Db 719 TCATCCCGTTTGTGATCATCTCTTGTACAGACTCTCATCCCAAACTTAAATCAA 778
 Qy 307 ySaspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrI 327
 Db 779 AAGATCNGATATGGCTGGGTACATCAAGCGCTCTCTCTCATCTTGTGAATTTCCACCA 838
 Qy 327 leCysPheAlaProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyHisA 347
 Db 839 TCTGCTTCCCTCCCAACCAAG----- 858
 Qy 347 snThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerC 367
 Db 858 ----- 858
 Qy 367 ysLeuAspProPheLeuTyrPheValMetSerLysValValAspGlnLeuAsnPro**S 387
 Db 858 ----- 858
 Qy 387 erAlaMetAlaArgProLeu***ArgProArgArgAspIleTrpGluAspIleHisAlaT 407
 Db 859 -----NNNNNGATATCTGGGAAGACGTACATGCTT 889
 Qy 407 xp 407
 Db 890 GG 891
 RESULT 10
 ID AAQ84559
 XX AAQ84559 standard; cDNA; 2732 BP.
 AC AAQ84559;
 XX
 XX 22-AUG-1995 (first entry)
 DE Murine C140 receptor cDNA.
 XX
 XX G-protein-coupled receptor; G-protein; C140 receptor; ss.
 OS Mus musculus.
 XX
 XX Key Location/Qualifiers
 XX CDS 73..1272
 FT /*tag= a

XX WO9503318-A.
 PN 02-FEB-1995.
 PD 26-JUL-1994; 94WO-US08536.
 PF 26-JUL-1993; 93US-0097938.
 PR (CORT-) COR THERAPEUTICS.
 PA Scarborough RM, Sundelin J;
 PI WPI; 1995-075182/10.
 XX P-PSDB; AAR66922.
 PT New DNA encoding recombinant C140 receptor - and novel agonists
 PT and antagonists, and specific antibodies with therapeutic and
 PT diagnostic applications.
 XX Claim 1; Fig 10; 57pp; English.
 XX A cDNA library from a mouse stomach was constructed in lambda gt10
 CC and screened with a probe encompassing the C140 genomic clone (see
 CC AAQ84557). A single phage clone was isolated and cut with EcoRI. The
 CC insert was cloned in pBluescript and pSG5 and sequenced. The
 CC complete nt sequence and deduced AA sequence is given in AAQ84559 &
 CC AAR66922. 5' RACE resulted in the addition of only 27 bps to the 5'
 CC end. The 5' end of the apparent coding region differs from the 5'
 CC end of the ORF of genomic DNA; it is believed that the 5' end of
 CC the cDNA sequence is correct.
 XX SQ Sequence 2732 BP; 650 A; 676 C; 633 G; 773 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,71e-44 Length: 2732
 Score: 584.00 Matches: 128
 Percent Similarity: 51.52% Conservative: 76
 Best Local Similarity: 32.32% Mismatches: 126
 Query Match: 27.34% Indels: 66
 DB: 16 Gaps: 10
 US-09-208-629F-3 (1-407) x AAQ84559 (1-2732)
 Qy 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuIleLeuValAlaAlaGlyLeu 27
 Db 61 CCACGTCGCGGATGCGAAGTCTCAGCTGGCGTGGCTGCTG-----GGAGGTATC 111
 Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
 Db 112 ACCCTTCTGGCGGCTCGGTCTCTCTCAGCGGACCGAGAACCTTGACCGGGACGCAAC 171
 Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
 Db 172 AACAGTAAAGGAAGAAGTCTTATTGCGAGATTAGAAACCCAGCCTCAATCACTGGGAAA 231
 Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGlyGly 73
 Db 232 GGGGTTCCGGTAGAACCGGCTTTTCCATCGATGAGTCTCTCGCTCCATCTCACCAGG 291
 Qy 74 TrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
 Db 291 ----- 291
 Qy 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerIleuSerThrGlnValIlePro 113
 Db 292 -----AAGCTGACCACGGTCTTTCTTCCG 315
 Qy 114 AlaIleTyrIleLeuLeuPheValGlyValProSerAsnIleValThrLeuTyrLys 133
 Db 316 GTGCTACATTATTGTTGTGTTGATGTTTCCAGTAATGCGATGCGCCCTCTGATC 375
 Qy 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152


```

Db      376  TTCTTTTCGAAAGAAAGAAACCCCGCGGTGATTTACATGGCCACCTGGCGCTTG 435
Qy      153  AlaAspLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
Db      436  GCGACCTCTCTCTGTCTGCTGTTCCCTCCCTGAAGATCTCTACCACTACATGGCAAC 495
Qy      173  AsnTyrValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
Db      496  AACTGGGTCTACGGGGAGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTATGATG 555
Qy      193  TyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
Db      556  TATTGCTCCATCTCTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615
Qy      213  ProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Db      616  CCCATGGGACACCCCGAGG---AAGAAAGGCAAAACATCCCGTCTCTCTTGGCAATC 672
Qy      233  TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeu 252
Db      673  TGGCTCTGATTTTCTGGTCACCATCCCTTTGTATGTCTATGAGCAGACCATCTACAT 732
Qy      253  ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Db      733  CCAGCATTTGAACATCACCATCTGTCCAGATGTGCTGCTGAG---GAGGTATTGGTGGG 789
Qy      273  PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Db      790  GACATGTTCAATTAATCTCTCTACTGCGCATTTGGAGTCTCTCTGTTCCCGGCCCTCT 849
Qy      293  IleIlePheCysTyrThrThrLeuIle-----ATCCGACTCATCATCACCGTGTG 957
Db      850  ACTGCATCTGCTGCTGCTCATGATCAAGCGCTCGCTCTCTGCTGATGGATGAACAC 909
Qy      303  LysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeu 322
Db      910  TCAGAGAAGAAAGGACAGAGGCT-----ATCCGACTCATCATCACCGTGTG 957
Qy      323  ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db      958  GCCATCTACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
Qy      343  TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db      1018  ATCAAAACCCAGGAGGAGCCAGCTGTACCCCTCTACCTGTGCGCCCTCTGCTGCTG 1077
Qy      363  SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db      1078  ACCCTCAACAGCTCATAGACCCCTTTGCTCTATTACTTTGCTCAAAA 1125

RESULT 11
ID      AAT32038
XX      AAT32038 standard; cDNA; 2732 BP.
AC      AAT32038;
XX      AAT32038;
DT      02-APR-1997 (first entry)
DE      Murine C140 receptor cDNA clone.
KW      C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW      antagonist; hypertension; hypotension; blood pressure; ss.
XX      Mus sp.
XX      Location/Qualifiers
FH      Key 73..1272
FT      CDS
FT      sig_peptide 73..165
FT      /tag= a
FT      /tag= b
FT      /note= "the signal sequence differs from a
FT      genomic DNA

```

```

FT      segment encoding this receptor (see AAT32036),
FT      the signal sequence given for this clone is
FT      believed to be the correct sequence"
FT      mat_peptide 166..1269
XX      /*tag= c
XX      WO9623225-A1.
XX      01-AUG-1996.
XX      25-JAN-1996; 96WO-0501179.
XX      25-JAN-1995; 95US-0390301.
XX      (CORT-) COR THERAPEUTICS INC.
XX      Scarborough RM, Sundelin J;
XX      WPI; 1996-362813/36.
XX      P-PSDB; AAW01954.
XX      Vector for expression C140 cell surface receptor in host cell
XX      useful to identify C140 agonist and antagonists, which are
XX      antihypertensives and elevators of blood pressure, respectively
XX      Example 4; Fig 10A-B; 60pp; English.
XX      AAT32038 encodes the murine C140 receptor (C140R). The sequence may be
XX      engineered so as to allow the recombinant expression of C140R in a
XX      suitable host cell, i.e. by removing the native expression-control
XX      sequence and replacing them with control sequences operable in the
XX      host. Such a recombinant receptor can be expressed on the surface of
XX      oocytes, this provides a good assay system for identifying
XX      agonists/antagonists of C140R. The C140 receptor is a G-protein linked
XX      receptor and a member of the "seven-pass" transmembrane receptor
XX      superfamily (peptide chain of the receptor passes through the cell
XX      membrane seven times, producing seven transmembrane regions within
XX      the receptor molecule). The C140 receptor is involved in controlling
XX      blood pressure. C140 antagonists (see AAW01942-W01951) are useful to
XX      inhibit signalling from this receptor, resulting in an increase in
XX      blood pressure and are therefore useful in pharmaceuticals for the
XX      treatment of hypotension (low blood pressure). Conversely agonists
XX      (see AAW01914-W01941) of C140 are useful in pharmaceuticals for the
XX      treatment of hypertension (high blood pressure).
XX      Sequence 2732 BP; 650 A; 676 C; 633 G; 773 T; 0 other;
SQ

```

Alignment Scores:

```

Pred. No.: 1.71e-44 Length: 2732
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservative: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 17 Gaps: 10

```

US-09-208-629F-3 (1-407) x AAT32038 (1-2732)

```

Qy      8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuLeuValAlaAlaGlyLeu 27
Db      61 CCACGTCGCGGGATGCGAAGATCTTCAGCCTGGCGTGGCTGCTG---GGAGGTATC 111
Qy      28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
Db      112 ACCCTTCTGGCGCCCTCGGTCTCTCTGACGCCGACCGAGAACCTTGCACCGGACGCAAC 171
Qy      44 AsnSer-----AlaLysProThrLeuThrLys 53
Db      172 AACAGTAAAGGAAGAAGTCTTATTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAA 231
Qy      54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
Db      232 GGGGTTCCGTTAGAACACGAGCTTTTCCATCGATGATGAGTCTCTCGGTCCATCTCACC 291

```


Qy 74 TrpThrGlyAlaThrThrThrIleLeuAlaGluCysProGluAspSerIleSerThrLeu 93
Db 291 ----- 291
Qy 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
Db 292 -----AAGCTGACCCAGCGTCTTCTTCGCG 315
Qy 114 AlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLys 133
Db 316 GTCGCTACATTATTGTGTGTGTTGATGTTGTCAGAGTATGCGATGCGCTCGATC 375
Qy 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
Db 376 TTCCTTTCCGAACGAGAGAAACACCCCGCGTGATTTACATGGCCAACTGGCGCTTG 435
Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
Db 436 GCGGACCTCTCTGTGTCATCTGTTCCCTGAGATCTCTACACCTACATGGCAAC 495
Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
Db 496 AACTGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATG 555
Qy 193 TyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
Db 556 TATTGCTCCATCTCTTACACCTGCTCAGCGTGCAGAGTACTGGGTGATCGTGAAC 615
Qy 213 ProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Db 616 CCATGGGACCCCGAGG---AAGAGGCAACATCGCGTGGCGTCTCTTGGCAATC 672
Qy 233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeu 252
Db 673 TGGCTCTGATTTTCTGGTCACCATCTCTTGTATGTATGAGAGACCATCTACATT 732
Qy 253 ValHisSerGluLeuThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Db 733 CCAGCAVTGAACATCACACCTGTCCAGATGTGCTGCTGAG---GAGGTATTGGTGGG 789
Qy 273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Db 790 GACATGTTCAATTACTTCTCTCTACCTGAGTATGGAGTCTTCTGTTCCCGGCCCTCTT 849
Qy 293 IleIlePheCysTyrThrThrLeuIle-----His 302
Db 850 ACTGCATCTCGCTACGGTCTCATGATCAAGACGCTCGCTCTTCTGCTATGGATGAAC 909
Qy 303 LysLeuLysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuIleLeu 322
Db 910 TCAGAGAAGAAAGGCGAGGGCT-----ATCGGACTCATCATCCGCTGCTG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db 958 GCCATGACTTCTATCTGCTTGTCTTGTAGCAACCTTCTGCTGAGTATGATTTCTCTA 1017
Qy 343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db 1018 ATCAAAACCCAGGAGGAGCCAGCTACGCCCTCTACCTTGTGCGCTCTGCTGCTG 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1078 ACCCTCAACAGCTGATAGACCCCTTTGTCTATTACTTTTCTCTCAAAA 1125
RESULT 12
AAQ84557
ID AAQ84557 standard; DNA; 1477 BP.
XX
XX AC
XX AAQ84557;
XX
DT 22-AUG-1995 (first entry)
XX
DE Murine C140 receptor genomic DNA.

XX G-protein-coupled receptor; G-protein; C140 receptor; ss.
XX Mus musculus.
OS
FH Key Location/Qualifiers
CDS 232..1188
FT /*tag= a
FT
XX WO9503318-A.
XX
XX 02-FEB-1995.
XX
XX 26-JUL-1994; 94WO-US08536.
XX
XX 26-JUL-1993; 93US-0097938.
XX
XX (CORT-) COR THERAPEUTICS.
XX
XX Scarborough RM, Sundelin J;
XX
XX WPI; 1995-075182/10.
XX
XX P-PSDB; AAR66920.
XX
XX New DNA encoding recombinant C140 receptor - and novel agonists
XX and antagonists and specific antibodies with therapeutic and
XX diagnostic applications.
XX
XX Disclosure; Fig 1; 57pp; English.
XX
XX A mouse cosmid genomic library (obtd. from Dr R.A. Wetsel, Washington
XX Univ. School of Medicine, St Louis, Missouri) was screened with two
XX 32P-labeled oligos corresp. to bp 190-249 and 742-801 of the bovine
XX substance K receptor cDNA. In one of the clones isolated (Cl40) the
XX hybridising region was localised to a 3.7 kb psi fragment. This
XX fragment was subcloned into pBluescript vector. The hybridising and
XX adjacent regions were sequenced. The nt sequence and the deduced AA
XX sequence are given in AAQ84557 & AAR6920 respectively. The published
XX DNA sequence appears to have been printed with the last two lines
XX in the wrong order. This has been corrected in this GeneSeq entry.
XX
SQ Sequence 1477 BP; 341 A; 414 C; 320 G; 402 T; 0 other;

Alignment Scores:
Pred. No.: 5,88e-44 Length: 1477
Score: 574.50 Matches: 110
Percent Similarity: 59.66% Conservative: 66
Best Local Similarity: 37.29% Mismatches: 102
Query Match: 26.90% Indels: 17
DB: 16 Gaps: 5

US-09-208-629f-3 (1-407) x AAQ84557 (1-1477)
Qy 95 ValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db 406 ATCGATGAGTCTCTCGTCCATCTCTACCGGAGAGTACACCGGTCTTCTTCGCGTC 465
Qy 115 IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
Db 466 GTCTACATTATTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 525
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 526 CTTTCCGAG 585
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 586 GACCTCCTCTGTCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 645
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db 646 TGGGTCTACGGGAGGCGCTGTGCAAGGTGCTCATTTGGCTTTTCTTATGTAACATGAT 705

```
Qy 194 CysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGCTCCATCTCTTCATGACCTGCTGCGTGCAGAGGTACTGGGTGATCGTGAACCC 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrp 233
Db 766 ATGGGACACCCAGG---AAGAAGGCAACATGCGCGTCTCTTGGCAATCTGG 822
Qy 234 ValMetValPheLeuTyrMetLeuPheValIleLeuLysGlnGluTyrHisLeuVal 253
Db 823 CTCCTGATTTTCGGTCACATCCCTTTGTATGTCAAGAGCAGACCATCTACATCCA 882
Qy 254 HisSerGluIleThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
Db 883 GCATTGAACATCACCATCTGCAGATGTGCTGCTGAG---GAGGTATTGGTGGGGAC 939
Qy 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuLeuProPheValIle 293
Db 940 ATGTTCAATTACTTCTCTCACTGGCCATTGGAGTCTTCTGTTCCGGCCCTCTTACT 999
Qy 294 IlePheCysTyrThrThrLeuIleHisLysLeuLysSer----- 306
Db 1000 GCATCTCCCTACGCTCATGATCAAGACGCTCGCTCTCTGCTATGGATGAACACTCA 1059
Qy 307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323
Db 1060 GAGAACAAAGGACAGGGCT-----ATCGACTCATCATCCGCTGCTGCC 1107
Qy 324 IlePheThrLeuCysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyr 343
Db 1108 ATGACTTCTCTCTTCTGCTCTAGCAACCTTCTGCTGCTAGTGCATTATTCTTAATC 1167
Qy 344 TyrTyrHisThrAspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySer 363
Db 1168 AAAACCCAGGAGGACCGCTACGCTTACCTTGTGCGCCCTCTGCTGCTGAC 1227
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1228 CTCACAGCTGCATAGACCCCTTGTCTATTAATCTTCTCTCAAAA 1272

RESULT 13
AAT32036
ID AAT32036 standard; DNA; 1477 BP.
XX
AC AAT32036;
XX
DT 01-APR-1997 (first entry)
XX
DE Murine C140 receptor-encoding DNA.
XX
KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 232..1419
FT sig_peptide /*tag= a
FT 232..312
FT /*tag= b
FT /*note= "putative signal sequence, differs from
FT signal sequence shown in a cDNA clone of
FT this receptor (see AAT32038), the signal
FT sequence given for the cDNA clone is
FT believed to be the correct sequence"
FT mat_peptide 313..1416
FT /*tag= c
XX
PN W09623225-A1.
XX
PD 01-AUG-1996.
XX
PF 25-JAN-1996; 96WO-US01179.
```

```
XX
PR 25-JAN-1995; 95US-0390301.
XX
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Scarborough RM, Sundelin J;
XX
DR WPI; 1996-362813/36.
XX
DR P-PSDB; AAW01952.
XX
PT Vector for expression C140 cell surface receptor in host cell
PT useful to identify C140 agonist and antagonists, which are
PT antihypertensives and elevators of blood pressure, respectively
XX
PS Example 1; Fig 1A-B; 60pp; English.
XX
CC AAT32036 encodes the murine C140 receptor (C140R), including a tentative
CC signal sequence. The sequence may be engineered so as to allow the
CC recombinant expression of C140R in a suitable host cell, i.e. by
CC removing the native expression-control sequence and replacing them with
CC control sequences operable in the host. Such a recombinant receptor
CC can be expressed on the surface of oocytes, this provides a good assay
CC system for identifying agonists/antagonists of C140R. The C140
CC receptor is a G-protein linked receptor and a member of the "seven-
CC pass" transmembrane receptor superfamily (peptide chain of the
CC seven transmembrane regions within the receptor molecule). The C140
CC receptor is involved in controlling blood pressure. C140 antagonists
CC (see AAW01942-W01951) are useful to inhibit signalling from this
CC receptor, resulting in an increase in blood pressure and are therefore
CC useful in pharmaceuticals for the treatment of hypotension (low blood
CC pressure). Conversely agonists (see AAW01914-W01941) of C140 are useful
CC in pharmaceuticals for the treatment of hypertension (high blood
CC pressure).
XX
SQ Sequence 1477 BP; 341 A; 415 C; 319 G; 402 T; 0 other;

Alignment Scores:
Pred. No.: 9,03e-44 Length: 1477
Score: 572.50 Matches: 110
Percent Similarity: 59.32% Conservatives: 65
Best Local Similarity: 37.29% Mismatches: 103
Query Match: 26.80% Indels: 17
DB: 17 Gaps: 5

US-09-208-629F-3 (1-407) x AAT32036 (1-1477)
Qy 95 ValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db 406 ATCGATGAGTCTCTGCTCCATCTCACGGGAAGCTGACCACGGTCTTCTTCGGTC 465
Qy 115 IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
Db 466 GTCTACATTATTGTGTGTGATTGTTGTTGCCAGTAATGGCATGCCCTCTCGATCTTC 525
Qy 135 SerLeuArgThr--LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 526 CTTTTCGGAACGAGAGAAACACCCCGCGTGATTTACATGGCCACCTTGGCTTGCC 585
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 586 GACCTCTCTCTCTCATCTGTTCCCTTCCCTGAGATCTCTTACCACCTACATGCAACAC 645
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db 646 TGGGTCTACGGGAGGCCCTGTGCAAGTGCTCATTTGGCTTTTCTATGTAACATGTAT 705
Qy 194 CysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGCTCCATCTCTTCTCATGACCTGCTGAGAGGTACTGGGTGATCGTGAACCC 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrp 233
```

Db 766 ATGGGACACCCAGG---AAGAAGGCAAAACATCGCCGTTGGCTCCTCCTTGGCAATCGG 822

Qy 234 ValMetValPheLeuTyMetLeuProPheValIleLeuLysGlnGluTyHisLeuVal 253

Db 823 CTCCTGATTTTCTGGTCACCATCCTTTGTATGTCATGAAGCAGCACCATCTACATTCGA 882

Qy 254 HisSerGluIleThrCysHisAspValValAspAlaCysGluSerProSerSerPhe 273

Db 883 GCATTGAACATCACCATCTGCAGATGTGCTGCTGAG---GAGGTATTGGTGGGGGAC 939

Qy 274 ArgPheTyTrpPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293

Db 940 ATGTTCAATTACTTCTCTACTCGCCATTGGAGTCTTCTGTTCCCGCCCTTCTTACT 999

Qy 294 IlePheCysTyTrpThrLeuIleHisLysLeuLysSer----- 306

Db 1000 GCATCTGCCTAGCGTCTCATGATCAAGACGCTCGCTCTTCTGCTATGGATGAACACTCA 1059

Qy 307 -----LysAspArgIleTrpLeuGlyTyTrpIleLysAlaValLeuLeuLeuVal 323

Db 1060 GAGAACAAAAGGCGAGAGGCT-----ATCGACTCATCATCAGCGTCTGCTGCC 1107

Qy 324 IlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyr 343

Db 1108 ATGACTTCATCTCTTCTGCTTAGCAACCTCTGCTGAGTGCATTATTCTCTAATC 1167

Qy 344 TyrTyHisAsnThrAspSerLeuTyTrpPheMetTyTrpLeuIleAlaLeuCysLeuGlySer 363

Db 1168 AAAACCCAGAGGAGAGCCAGCTCTAGCCCTCTACCTTGTGCGCCTCTGCTGTGACC 1227

Qy 364 LeuAsnSerCysLeuAspProPheLeuTyTrpPheValMetSerLys 378

Db 1228 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTTGTCTCAAAA 1272

RESULT 14

AAQ84558
ID AAQ84558 standard; DNA; 1255 BP.

XX AC AAQ84558;

XX DT 22-AUG-1995 (first entry)

XX XX Human C140 receptor genomic DNA.

XX KW G-protein-coupled receptor; G-protein; C140 receptor; ss.

XX OS Homo sapiens.

Key Location/Qualifiers
CDS 56..1197

FT FT /*tag= a

XX PN WO9503318-A.

XX XX 02-FEB-1995.

XX PF 26-JUL-1994; 94WO-US08536.

XX PR 26-JUL-1993; 93US-0097938.

XX PA (CORT-) COR THERAPEUTICS.

XX XX Scarborough RM, Sundelin J;

XX DR WPI; 1995-075182/10.

XX DR P-PSDB; AAR66921.

XX XX New DNA encoding recombinant C140 receptor - and novel agonists
PT and antagonists and specific antibodies with therapeutic and
PT diagnostic applications.

XX PS Disclosure; Fig 2; 57pp; English.

XX XX

CC The availability of genomic DNA encoding the mouse protease
CC C140 receptor (see Q84557) permitted the retrieval of the corresp.
CC human gene. A human genomic library cloned in the vector EMBL3 was
CC screened using the entire coding region of the murine clone as a
CC probe. The recovered human gene including the DNA sequence and the
CC deduced AA sequence are shown in Q84558 & R66921. Subsequent
CC experiments indicated that the human C140 gene is located in the
CC same region of the long arm of chromosome number 5 (5q12-5q13)
CC as has been reported for the human thrombin receptor gene.

XX SQ Sequence 1255 BP; 294 A; 320 C; 260 G; 381 T; 0 other;

Alignment Scores:

Pred. No.: 5,61e-43 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 16 Gaps: 7

US-09-208-629f-3 (1-407) x AAQ84558 (1-1255)

Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68

Db 116 ACCCTGTCTCTCTTCTTGTACAGGAACAATAGATCTCTTAAGGAAGAACGCTTATT 175

Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85

Db 176 GGTAAAGTTGATGGACATCCACGTCACCTGGAAAGAGGATTACAGTTGAA----- 226

Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyTrpLeuArgSer 105

Db 227 -----ACAGTCTTTTCTGTGGATGAGTTTCTGCATCTGCTCTCACTGGA 271

Qy 106 SerLeuSerThrGlnValIleProAlaIleTyTrpIleLeuLeuPheValGlyValPro 125

Db 272 AAATCAGCCATGCTCTTCTTCCAAATGTCTACAAATGTGTGTGGTGGGTTGCCA 331

Qy 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144

Db 332 AGTAACGGCATGGCCCTGTGGGTCTTTCTTCCGAACCTAAGAAAGAACCCCTGCTGTG 391

Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164

Db 392 ATTTACATGGCCAAATCTGGCTTGGTGACCTCTCTCTGTCTGTCTGTTCCCTTGAAG 451

Qy 165 IleAlaTyHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184

Db 452 ATTGCTATCACATACATACGCAACAACTGGATTATGGGGAGCTCTTTGTAATGTGCTT 511

Qy 185 ThrValValPheTyTrpGlyAsnMetTyTrpCysAlaIleLeuLeuThrCysMetGlyIle 204

Db 512 ATTGCGCTTTTCTATGGCAACATGTAATCTTCTTCCATCTCTTTCATGACCTGCCTCAGTGTG 571

Qy 205 AsnArgTyTrpLeuAlaThrAlaHisProPheThrTyTrpGlnLysLeuProLysArgSerPhe 224

Db 572 CAGAGGTATTGGGTCTATCGTAACCCCATGGGCACTCCAGG--AAGAAGGCAAAACATT 628

Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyTrpMetLeuProPheVal 244

Db 629 GCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTCTGCTGCACCATCTTCTTGTAT 688

Qy 245 IleLeuLysGlnGluTyHisLeuValHisSerGluIleThrThrCysHisAspValVal 264

Db 689 GTCGTAGAGCAGACCATCTTCTCTGCGCTGAACATCAGACCTGTGATGATGTTTG 748

Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyTrpPheValSer 280

Db 749 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCTCT 793

Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyTrpThrLeu 300

Db 794 CTGGCAATTGGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 853

```

QY 301 ILeHisLysLeuYsSer-----LysAspArgIleTrpLeuGlyTyr 314
DB 854 ATCAGATGTCGGATCTTCTGCCATGGATGAAACTCAGAGAGAAAGAGAGGCC 913
QY 315 ILeYsAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
DB 914 ATCAAACTCATGTGTCACCTGCTGGCCATGTACCTGATCTGCTCACTCTAGTAACCTT 973
QY 335 ILeLeuValIleHisAlaAsnTyrTyrTrpHisAsnThrAspSerLeuTyrPheMet 354
DB 974 CTGCTTGTGGTCAATTATTTCTGATTAGAGCCAGGCCAGCCATGCTATGCCCTG 1033
QY 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
DB 1034 TACATTGAGCCCTCTGCCCTCTTACCCTTAACAGCTGCATCGACCCCTTGTCTATTAC 1093
QY 375 ValMetSer 377
DB 1094 TTTGTTTCA 1102
RESULT 15
AAT32037
ID AAT32037 standard; DNA; 1255 BP.
XX
AC AAT32037;
XX
DT 01-APR-1997 (first entry)
XX
DE Human C140 receptor-encoding DNA.
XX
KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 56..1197
FT /tag= a
FT sig_peptide 56..136
FT /tag= b
FT /note= "putative signal sequence, differs from
FT signal sequence shown in a cDNA clone of
FT this receptor (see AAT32039 - this clone also
FT encodes a mature receptor having four amino
FT acid changes, when compared to this sequence)"
FT mat_peptide 137..1194
FT /tag= c.
XX
XX WO9623225-A1.
XX
XX 01-AUG-1996.
XX
XX 25-JAN-1996; 96WO-US01179.
XX
XX 25-JAN-1995; 95US-0390301.
XX
XX (CORT-) COR THERAPEUTICS INC.
XX
XX ScARBorough RM, Sundelin J;
XX WPI; 1996-362813/36.
XX
XX P-PSDB; AAW01953.
XX
XX Vector for expression C140 cell surface receptor in host cell
XX useful to identify C140 agonist and antagonists, which are
XX antihypertensives and elevators of blood pressure, respectively
XX
XX Example 2; Fig 2A-B; 60pp; English.
XX
XX AAT32037 encodes the human C140 receptor (C140R), including a tentative
XX signal sequence. The sequence may be engineered so as to allow the
XX recombinant expression of C140R in a suitable host cell, i.e. by

```

```

CC removing the native expression-control sequence and replacing them with
CC control sequences operable in the host. Such a recombinant receptor
CC can be expressed on the surface of oocytes, this provides a good assay
CC system for identifying agonists/antagonists of C140R. The C140
CC receptor is a G-protein linked receptor and a member of the "seven-
CC pass" transmembrane receptor superfamily (peptide chain of the
CC receptor passes through the cell membrane seven times, producing
CC seven transmembrane regions within the receptor molecule). The C140
CC receptor is involved in controlling blood pressure. C140 antagonists
CC (see AAW01942-W01951) are useful to inhibit signalling from this
CC receptor, resulting in an increase in blood pressure and are therefore
CC useful in pharmaceuticals for the treatment of hypotension (low blood
CC pressure). Conversely agonists (see AAW01914-W01941) of C140 are useful
CC in pharmaceuticals for the treatment of hypertension (high blood
CC pressure).
XX
SQ Sequence 1255 BP; 294 A; 320 C; 260 G; 381 T; 0 other;

Alignment Scores:
Pred. No.: 5,61e-43 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 17 Gaps: 7

US-09-208-629f-3 (1-407) x AAT32037 (1-1255)
QY 49 ThrLeuThrIleYsSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
DB 116 ACCCTTGTCTTCTTCTTGTACAGGACCAATAGATCTCTAAAGGAAGACCTTATT 175
QY 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleYsAlaGluCys 85
DB 176 GGTAAAGTGTAGTGGCACATCCACGTCCTGCTAAAGAGAGGTATACATTTGAA----- 226
QY 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
DB 227 -----ACAGTCITTTCTGTGGATGAGTTCCTCATCTGCTCTCACTGGA 271
QY 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
DB 272 AAACCTGACCACCTGCTCTTCTTCCAAATGTCTACACAAATGTGTGTGTGGTGGTGGCCA 331
QY 126 SerAsnIleValThrLeuTrpYsLeuSerLeuArgThr---LysSerIleSerLeuVal 144
DB 332 AGTAACGGCATGCCCTGTGGGTCTTTCTTTCCGAACCTAAGAAAGAACCCCTGCTGTG 391
QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
DB 392 ATTTACATGGCCAAATCTGGCTTGGCTGACCTCTCTCTCATCTGGTTCCCTTGAAG 451
QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
DB 452 ATTGCCTATCACATACATACGCAACAACTGGATTATATGGGAAGCTCTTTGTATATGCTT 511
QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIle 204
DB 512 ATGGGCTTTTCTATGGCAACATGATCTTTCATTTCTTATGACCTGCTCAGTGTG 571
QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnYsLeuProYsArgSerPhe 224
DB 572 CAGAGGTATGGGTATCGTGAACCCCATGGGCACTCCAGG---AAGAAGGCAACATT 628
QY 225 SerLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPheVal 244
DB 629 GCCATTGGCATCTCCCTGGCAATATGCTGCTGATTCGTGCTCACCCTCCCTTGTAT 688
QY 245 IleLeuYsGlnGluThrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
DB 689 GTGCTGAAGCAACACCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 748
QY 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280

```

Db 749 -----CTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCT 793
 Qy 281 LeuAlaPheGlyLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
 Db 794 CTGGCCATGGGGTCTTTCTGTTCCAGCCTTCTCTACAGCCTCTGCTATGCTGATG 853
 Qy 301 IleHisTysLeuSer-----LysAspArgIleTrpLeuGlyTyr 314
 Db 854 ATCAGAATGTCGATCTTCTGCCATGATCAAACTCAGAGAAGAAAGAGAGGCC 913
 Qy 315 IleIysAlaValLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 914 ATCAACTCATGTGCTGCTGGCCATGATCTGCTTCACTCTCTAGTAACCTT 973
 Qy 335 IleLeuValIleHisHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
 Db 974 CTGCTCTGGTGCAATTATTTCTGATTAGAGCCAGGCCGATGCTATGCCCTG 1033
 Qy 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPropheLeuTyrPhe 374
 Db 1034 TACATTGAGCCCTCTGCTCTCTACCCCTTAACAGCTGCATCGACCCCTTGTCTATTAC 1093
 Qy 375 ValMetSer .377
 Db 1094 TTTGTTTCA 1102

RESULT 16

ID AAZ50773 standard; DNA; 1289 BP.
 AC AAZ50773;

DT 31-MAY-2000 (first entry)
 DE Human protease activated receptor-2 DNA.

KW Human; PAR-2; protease activated receptor;
 KW antisenase molecule; PAR antibody; cytostatic; therapeutic;
 KW metastatic tumour cell; placental implantation; invasive cell; ds.

OS Homo sapiens.

PN W0200008150-A1.

PD 17-FEB-2000.

PF 05-FEB-1999; 99WO-IL00079.

PR 07-AUG-1998; 98IL-0125698.

PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.

PI Bar-Shavit R;

DR WPI; 2000-205706/18.

XX Treating metastatic tumor cells useful for treating disorders involving
 XX placenta implantation in a female comprises administration of an
 XX antisenase molecule complementary to an RNA sequence of a protease
 XX activated receptor protein -

XX Example 3; Fig 9; 46pp; English.

XX The patent discloses a method to treat metastatic tumour cells using
 XX an antisenase molecule comprising a polynucleotide complementary to an
 XX RNA sequence of a protease activated receptor (PAR) protein, or an
 XX antibody capable of binding to a PAR protein. The antisenase molecules and
 XX antibodies of PAR protein are also used to treat disorders associated
 XX with implantation of placenta. The present sequence is a human PAR-2
 XX DNA used for producing antisenase molecules for treating invasive
 XX cells.

XX

SQ Sequence 1289 BP; 300 A; 320 C; 270 G; 399 T; 0 other;

Alignment Scores:

Pred. No.: 5-81e-43 Length: 1289
 Score: 563.00 Matches: 115
 Percent Similarity: 54.52% Conservative: 72
 Best Local Similarity: 33.53% Mismatches: 128
 Query Match: 26.36% Indels: 28
 DB: 21 Gaps: 7

US-09-208-629F-3 (1-407) x AAZ50773 (1-1289)

Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
 Db 72 ACCCTTGTCTTCTTCTTGTACAGAACCAATAGATCTCTAAAGAGAGAGACCTTATT 131
 Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleIleAlaGluCys 85
 Db 132 GGTAAGGTGTATGGGACATCCCACTGCTGGAAGAGAGAGTTACAGTTGAA----- 182
 Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
 Db 183 -----ACAGTCTTTCTGTGGATGAGTTTCTGCATCTGTCTCTCACTGGA 227
 Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
 Db 228 AAACAGCACCATGCTCTTCTTCCAAATGTCTACAAATGTGTGTGGTGGTGGTGGCA 287
 Qy 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
 Db 288 AGTAACGGCATGGCCCTGTGGGTCTTTCTTTCGAACTAAGAGAGAGACCCCTGTGTG 347
 Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
 Db 348 ATTACATGGCCAACTGCGCTTGGCTGACCTCTCTGTCTCATCTGGTTCCTCTCAAG 407
 Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
 Db 408 ATTGCTATACATACATACATGCAACAACTGATTTATGGGAAGCTCTTTGTAATGCTT 467
 Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
 Db 468 ATTGGCTTTTCTATGGCAACATGATGTCTTCATTTCTTTCATGACCTGCCTCAGTGTG 527
 Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
 Db 528 CAGAGGTATGGGTCTATCGTAACCCCATGGGCACTCCAGG---AAGAAGGCAACATT 584
 Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
 Db 585 GCCATTGGCATCTCCCTGGCAATATGGTCTGATTTCTGTGTGTCAACATCCCTTTGTAT 644
 Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
 Db 645 GTCTGAAGCAGACCATCTTCATCTCCCTGCAACATCACGACCTGTCTATGATGTTTG 704
 Qy 265 AspAlaCysGluSerProSerPhe-----ArgPheTyrTyrPheValSer 280
 Db 705 -----CCTGAGCAGCTTGTGGGGAGACATGTTCAATTACTTCTCTCT 749
 Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
 Db 750 CTGGCCATTGGGTCTTCTTCTTCCAGCCTTCTCTCACAGCCTCTGCCTATGCTGATG 809
 Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
 Db 810 ATCAGAATGCTGGCATCTTCTGCAATGATGAACCTCAGAGAAGAAAGAGAGAGGCC 869
 Qy 315 IleIysAlaValLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 870 ATCAAACTCATGTGCTACTGTCTGGCCATGCTACTGATCTGCTTCACTCTCTAGTACCTT 929
 Qy 335 IleLeuValIleHisHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet 354

251 GTCCCTCGCTGGAAAACTGACCACTGCTTCCTTCCTCAATGTGCTACACAATGTGTTGTTGGC 310
122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSer 140
311 GTGGGTTTGCCAAAGTAACGGCATGGCCCTATGGTCTTCTTTTCCGAACTAAGAAGAAG 370
141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
371 CACCCTCGCTGATTTACATGGCCAACTCTGGCCCTTGCGCTGCACCTCTCTCTCATCTGG 430
161 LeuProPheLysIleAlaTrpHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
431 TTCCCTCTGAAGATTCCTATCATACATGAGCACTGGATTTATGGGAAGCTCT 490
181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
491 TGTAAATGCTTATTGGCTTTTTCATATCGCAACATGFACTGTTCCTATCTCTTCATGACC 550
201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
551 TGCCCTCAGTGTGAGAGGATTTGGGTCTATCGTGAACCCCATCGGGCACTCCAGG---AAG 607
221 LysArgSer:PheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
608 AAGCAAAATTCGCCAATGGCACTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCAACC 667
241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
668 ATCCCTTTGTATGCTGTGAAGCAGACCATCTTCATTCCTGCCCTGAAACATCAGCACCTGT 727
261 HisAspValValAspAlaCysGluSerProSerSer:Phe-----ArgPheTyr 276
728 CATGATGTTTG-----CCTGAGCAGCTCTTGGTGGGAGACATGTTTCAAT 772
277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCys 296
773 TACTTCCTCTCTCTGCCCATTTGGGGTCTTCTGTTCCAGCCTTCCTCAGACGCTCTGCC 832
297 TyrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
833 TATGTCTGATGATCAGATGCTGCCATCTTCTGCCATGGATGAAAACTCAGAGAAGAAA 892
311 TrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
893 AGGAAGAGGGCCCATCAAACTCATTTGTCCTGGCATGTACCTGATCTGCTTCACT 952
331 ProThrAsnIleIleValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
953 CCTAGTAACCTCTGCTGTGGTGCAATATTTCTGATTAAAGCCAGGCGCCAGAGCCAT 1012
351 LeuTyr:PheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
1013 GTCTATGCCCTGTACATTTGAGCCCTCTGCCTCTCTACCCCTTAACAGCTGCATCGACCCC 1072
371 PheLeuTyr:PheValMetSer 377
1073 TTTGCTATTACTTTGTTTCA 1093

RESULT 19
AAT32039
ID AAT32039 standard; cDNA; 1414 BP.
XX
XX AAT32039;
XX
XX 02-APR-1997 (first entry)
XX Human C140 receptor cDNA clone.
XX
XX C140 receptor; G-protein linked; coupled; seven pass; agonist;
XX antagonist; hypertension; hypotension; blood pressure; ss.
XX
XX Homo sapiens.
XX

FH Key Location/Qualifiers
 FT CDS 50..1243
 FT /*tag= a
 FT sig_peptide 50..130
 FT /*tag= b
 FT /note= "the signal sequence, differs from the
 FT signal sequence shown in a genomic clone of
 FT this receptor (see AAT32037). The cDNA clone
 FT also encodes a mature receptor having four
 FT amino acid changes, when compared to the
 FT genomic clone"
 FT mat_peptide 131..1240
 FT /*tag= c
 PN W05623225-A1.
 XX 01-AUG-1996.
 XX 25-JAN-1996; 96WO-US01179.
 XX 25-JAN-1995; 95US-0390301.
 XX (CORT-) COR THERAPEUTICS INC.
 XX Scarborough RM, Sundelin J;
 XX WPI; 1996-362813/36.
 DR P-PSDB; AAW01955.
 XX
 XX Vector for expression C140 cell surface receptor in host cell -
 FT useful to identify C140 agonist and antagonists, which are
 FT antihypertensives and elevators of blood pressure, respectively
 XX
 XX Example 5; Fig 11A-B; 60pp; English.
 XX
 CC AAT32039 encodes the human C140 receptor (C140R). The sequence may be
 CC engineered so as to allow the recombinant expression of C140R in a
 CC suitable host cell, i.e. by removing the native expression-control
 CC sequence and replacing them with control sequences operable in the
 CC host. Such a recombinant receptor can be expressed on the surface of
 CC oocytes, this provides a good assay system for identifying
 CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked
 CC receptor and a member of the "seven-pass" transmembrane receptor
 CC superfamily (peptide chain of the receptor passes through the cell
 CC membrane seven times, producing seven transmembrane regions within the
 CC receptor molecule). The C140 receptor is involved in controlling blood
 CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
 CC signalling from this receptor, resulting in an increase in blood
 CC pressure and are therefore useful in pharmaceuticals for the treatment
 CC of hypotension (low blood pressure). Conversely agonists (see AAW01914-
 CC AAW01941) of C140 are useful in pharmaceuticals for the treatment of
 CC hypertension (high blood pressure).
 XX
 SQ Sequence 1414 BP; 335 A; 361 C; 309 G; 409 T; 0 other;
 Alignment Scores:
 Pred. No.: 8,61e-42 Length: 1414
 Score: 551.00 Matches: 118
 Percent Similarity: 51.77% Conservative: 72
 Best Local Similarity: 32.15% Mismatches: 139
 Query Match: 25.80% Indels: 38
 DB: 17 Gaps: 8
 US-09-208-629F-3 (1-407) x AAT32039 (1-1414)
 Qy 22 LeuValAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
 Db 74 CTGCTGGGGCGCCGCACTCTCTCTAGCAGCCTCTCTCTGCTGAGCCATCCCAAGGA 133
 Qy 42 SerAspAsnSerAlaGlyLeuLeuPheProThrLeuThrIleLysSerPheAsnGlyProGlnAsn 61
 Db 134 ACCAATAGATCCCTTAAAGGAAGAACGCCCTTATTGGTAGGTTGATGCG----- 181

QY 62 ThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrIle 81
 DB 182 -----ACATCCCAAGTTCACAGTT 214
 QY 82 LysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGly 101
 DB 215 GAA-----ACAGCTTTTCTGTGGATGAGTTTCTGCATCT 250
 QY 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 121
 DB 251 GTCTCGCTGGAAACTGACCACTGTCTCTCCATTCACATTTGTCTACAAATGGTTGG 310
 QY 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSer 140
 DB 311 GTGGGTTTGCACAGTAACGGCATGCGCTTATGGTCTTTCTTCGGAACCTAAGAAGAAG 370
 QY 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
 DB 371 CACCCCTGTGTGATTTACATGGCCAATCTGGCTTGGCTGACCTCTCTCTCATCTGG 430
 QY 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
 DB 431 TTCCCTTGAAGATTGCCCTATCACATACATGCAACACTGGATTTATGGGAAGCTCT 490
 QY 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuThr 200
 DB 491 TGTAAATGCTTATTGGCTTTTCTATCGCAACATGACTGTTCATTTCTTCATGACC 550
 QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
 DB 551 TGCTCAGTGTGCAGAGGATTGGGTGTCATCGTGAACCCCATGGGCACCTCCAGG---AAG 607
 QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
 DB 608 AAGGCAAAATTCGCCATTCCTCCCTGGCAATATGCTGCTGACTCTGTGTGTCACC 667
 QY 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCys 260
 DB 668 ATCCCTTTGATGTCGTGAAGCAGACCATCTTCATTCCTGCTGCAACATCAGACCTGT 727
 QY 261 HisAspValValAspAlaCysGluSerProSerSerPhe-----ArgPheTyr 276
 DB 728 CATGATGTTTG-----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772
 QY 277 TyrPheValSerLeuAlaPhePheGlyPheLeuLeuProPheValIleIlePheCys 296
 DB 773 TACTTCTCTCTGTCGCCATTGGGTCTTTTGTTCCTCCAGCCTTCTCAGCCTCTGCC 832
 QY 297 TyrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
 DB 833 TATGCTGATGATCAGAAATGCTGCGATCTTCTGCCATGATGATAAACTCAGAGAAGAA 892
 QY 311 TrpLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAla 330
 DB 893 AGGAAGAGGGCCATCAAACTCATTTGTCACTGCTCTGGCATGTATCTGCTTCACT 952
 QY 331 ProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
 DB 953 CTTAGTAACCTTCTGCTTGTGGTCATTATTTCTGATTAAAGCCAGGCCAGGCCAT 1012
 QY 351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
 DB 1013 GTCTATGCCCTGTACATTGTAGCCCTCTGCTCTTACCTTTAACAGCTGCATCGACCCC 1072
 QY 371 PheLeuTyrPheValMetSer 377
 DB 1073 TTGTCTATTACTTTGTTCA 1093
 RESULT 20
 AAT93367
 ID AAT93367 standard; cDNA; 1312 BP.
 XX
 AC AAT93367;

XX 01-MAY-1998 (first entry)
 XX Mouse thrombin receptor cDNA.
 XX Transgenic animal; thrombin receptor; animal model; restenosis;
 KW knockout animal; antagonist; agonist; mouse; murine;
 KW Alzheimer's disease; wound repair; ss.
 XX Mus sp.
 XX Key Location/Qualifiers
 FH complement (612..631)
 FT /*tag= a
 FT /label= MTR-3
 FT 1294..1311
 FT /*tag= b
 FT /label= MTR-2
 XX W09740138-A1.
 XX 30-OCT-1997.
 XX 22-APR-1997; 97WO-US06575.
 XX 25-JUN-1996; 96US-0020544.
 XX 23-APR-1996; 96US-0016070.
 XX (ORTH) ORTHO PHARM CORP.
 XX Andrade-Gordon P, Leung W, Zhou L;
 XX WPI; 1997-535823/49.
 XX Transgenic animals with altered form of thrombin receptor gene -
 PT useful in study of receptor function and identification of
 PT antagonists and agonists for disease therapy
 XX Example 1; Fig 1; 47pp; English.
 CC The present sequence was used in the preparation of a transgenic
 CC non-human vertebrate with somatic and germ cells containing an
 CC altered form of a thrombin receptor (TR) gene, where the altered
 CC gene has replaced a wild type TR gene in the animal or an ancestor
 CC at an embryonic stage using embryonic stem cells. The transgenic
 CC animal provides an animal model to understand receptor function in,
 CC e.g. platelets, endothelial cells and neurons, and evaluate drug
 CC therapies modulating TR function/expression in human cells.
 CC Alterations rendering the wild type gene non-functional produce
 CC knockout animals, useful as non-human whole animal models for
 CC diseases involving TR equivalents in humans. Alternatively,
 CC substituting the naturally occurring gene for a gene from a 2nd
 CC species or with a mutation, results in animals producing the 2nd
 CC species or mutated gene products. These transgenic animals are
 CC useful for drug antagonist and agonist studies, creation of animal
 CC models of human diseases and investigation of disorders associated
 CC with human TR-mediated responses. Antagonists may be used
 CC therapeutically for cardiovascular problems, e.g. restenosis, to
 CC limit platelet adhesion during interventional procedures or to
 CC generate therapies for bone degenerative and neurodegenerative
 CC diseases, e.g. Alzheimer's. Agonists may be useful in wound repair.
 XX SQ Sequence 1312 BP; 247 A; 406 C; 328 G; 331 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.97e-38 Length: 1312
 Score: 514.50 Matches: 120
 Percent Similarity: 51.28% Conservative: 80
 Best Local Similarity: 30.77% Mismatches: 143
 Query Match: 24.09% Indels: 47
 DB: 18 Gaps: 13
 US-09-208-629F-3 (1-407) x AAT93367 (1-1312)

QY 20 LeuileLeuValAlaAlaGly-----LeuLeuPheLeuProValThr 33
 |||:|||||
 Db 30 TTGCTGATCGTCGCTCGCTCAGCCTGTGGGTCCCTTGCTGTCTTCCCGCGTCCCT 89
 QY 34 ValCysGlnSerGlyIleAsnValSerAspAsnSerAlaIysProThrLeuThrIleLys 53
 |||:|||||
 Db 90 ATGAGCCAGCCAGAAATCAGAGAGGAGAGATGCTACGGTGAACCCC-----CGC 137
 QY 54 SerPhe-----AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIle 71
 |||:|||||
 Db 138 TCATTCTTTCTAAGGAATCCAGTGAAATATACATTTGAACCTGGTCCCTTGGGGGATGAG 197
 QY 72 GluGly-----TrpThrGlyAlaThrThrThrIleLysAla 83
 |||:|||||
 Db 198 GAGGAGGAGGAGAAAAATGAAGCGCTCTGCTGGAGGGTGGGCGAGTCTACTTAAATATA 257
 QY 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
 |||:|||||
 Db 258 AGCTCCCTCCTCACACGCGCTCTCCCTTCATCTCCGAGGAGCGCTCCGATATCTG 317
 QY 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
 |||:|||||
 Db 318 ACCAGCCCTGGCTGAGCTCTTCATGCCCTCCGTGTACAGATTGTGTTTCATGTTCAGC 377
 QY 124 ValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLys---SerIleSer 142
 |||:|||||
 Db 378 CTTCTCTGAACGTCCTGCCATCGCAGTGTTCTGCTTGAGGATGAAGGTCAAGAAGCCG 437
 QY 143 LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuPro 162
 |||:|||||
 Db 438 GCGTGGTGTACATGCTGCACCTGGCCATGGCGGACGCTCTCTCGTGTGGTGTCTCCCT 497
 QY 163 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArg 182
 |||:|||||
 Db 498 TCACAGATCAGCTACTACTTCTCGGCACTGATTGGCAGTTTCGGGTCTGGAATGTGCCGT 557
 QY 183 IleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMet 202
 |||:|||||
 Db 558 TTGCGCACCGCAGCGTTTACTATAACATGTAGCCCTCCATCATGTCATGACGGTCATA 617
 QY 203 GlyIleAsnArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeu 219
 |||:|||||
 Db 618 AGCATTTAGCGGTTCTCTGGCGGTGGTGTATCCGATCCAGTCCCTGCTGGCGCACTCTG 677
 QY 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr 239
 |||:|||||
 Db 678 GCGGAGCGCAACTTCACT-----TGGCGTGTTCATTGGGTGATGGCCATCATGGGG 728
 QY 240 MetLeuProPheValIleLeuLysGlnIuTyrHisLeuValHisSerGluIleThrThr 259
 |||:|||||
 Db 729 GTGGTCCCTTCTCTCAAGGAGCAGACCCCGAGTTCGGGACTCAACATCACCACC 788
 QY 260 CysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheVal 279
 |||:|||||
 Db 789 TGGCAGCATCTCTCTAGTGAG---AACCTGATGCAAGGCTTTTACTCGTACTACTTCTCG 845
 QY 280 SerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThr 299
 |||:|||||
 Db 846 GCCTTCTCCGCCATCTTCTTCTGTGGCGTGTATCGTTCCACGGTCTGTCTACAGTCC 905
 QY 300 LeuIleHisLysLeuLysSer-----LysAspArgIleTrp 311
 |||:|||||
 Db 906 ATCATCCGGTGGCTGAGCTCTCCGGTTCGCCAGCGGACCAAGATCGCGGGCTTTC 965
 QY 312 LeuGlyTyrIleLysAlaValLeuLeuLeuIleLeuValIlePheThrIleCysPheAlaPro 331
 |||:|||||
 Db 966 -----TTCTGTCTCCCGG-----GTGTTCTGCATCTTTCATCTGCTGTGGGCC 1013
 QY 332 ThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyr-----HisAsn 347
 |||:|||||
 Db 1014 ACCACGTCCTCTGATTGTGCAC-----TACCTTTTCTCTCCGACAGTCTCTGGT 1064

Qy 348 ThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 367
 Db 1065 ACGGAGCAGCGTACTTGTGCTTACCTCTCTGCGTCTGTGTGACGAGCGTGAGTGTGCTGC 1124

Qy 368 LeuAspProPheLeuTyrPheValMetSer 377
 Db 1125 ATCGATCGTGTGATTTACTACTACGCGCTCC 1154

RESULT 21
 AAQ73590
 ID AAQ73590 standard; DNA; 1764 BP.
 AC AAQ73590;
 AC AAQ73590;
 DT 25-JUN-1995 (first entry)
 XX Fragment of the human thrombin receptor gene.
 DE XX
 DE XX
 KW TR; expression; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 376..1653
 FT /*tag= a
 XX
 PN WO9421789-A.
 XX
 PD 29-SEP-1994.
 XX
 PF 28-FEB-1994; 94WO-US02388.
 XX
 PR 25-MAR-1993; 93US-0038662.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Betlach MC, Turner GJ;
 XX
 DR WPI; 1994-317010/39.
 DR P-PSDB; AAR60698.
 XX
 PT Expression of heterologous proteins in halo-bacteria - using
 PT regulatory and stop sequences from halo-bacteria, pref. the
 PT bacterio-rhodopsin gene.
 XX
 PS Disclosure; Fig 14; 118pp; English.
 XX
 CC The sequence is that of the human thrombin receptor gene fragment.
 CC The gene is used to exemplify a new expression vector for producing
 CC heterologous polypeptides in a halobacterial host.
 CC See also AAQ73586-92.
 XX
 SQ Sequence 1764 BP; 382 A; 492 C; 421 G; 469 T; 0 other;

Alignment Scores:
 Pred. No.: 5,53e-38 Length: 1764
 Score: 511.50 Matches: 122
 Percent Similarity: 52.07% Conservative: 79
 Best Local Similarity: 31.61% Mismatches: 148
 Query Match: 23.95% Indels: 37
 DB: 15 Gaps: 14

US-09-208-629f-3 (1-407) x AAQ73590 (1-1764)

Qy 19 IleIleuLeuAlaAlaGlyLeuLeuPheLeuProValThr-----ValCys 35
 Db 355 ATCTCGTGTAGTACTGTGCTGATGTTGAGTTATTGCCAACAGCAGTGAGGGGTATCG 414
 Qy 36 GlnSerGlyIleAsnVal-----SerAspAsnSerAlaLysProThrLeuThr 51
 Db 415 CAGGCCCNAGTCCAGGGCGCTGGACTACAGGACCATGATGACGTCGACGCCACCTTATGAT 474
 Qy 52 IleLysSerPhe-----AsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAsp 70

RESULT 22
 AA232191

Db 475 CCCCGGTCTATTTCTTCTCAGGAACCCCAATGATAATATCAACATTTTGGGAGGATGAG 534
 Qy 71 IleGluGlyTrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysPro 86
 Db 535 GAGAAAAATGAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGCACTCCT 594
 Qy 87 GluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSer 106
 Db 595 CTTCAAAAAACAATCTCTCATTCATCTCAGAGATGCTCCCGGATATTTGACCACTCC 654
 Qy 107 LeuSerThrGlnValIleProIleTyrIleLeuLeuPheValValGlyValProSer 126
 Db 655 TGGCTCACACTCTTTTGTCCCATCTGTGTACACGGAGTGTGTGTAGTCAGCCTCCCACTA 714
 Qy 127 AsnIle-----ValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuVal 144
 Db 715 AACATCATGCCATCTGTTGTTCATCTCGAAATGAAGGTCAAGAG---CGGCGGGTG 771
 Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
 Db 772 GTGTACATGCTGCACCTGCCACGCGAGATGTCTGTTGTGTCTGTCTCCCTTTAAG 831
 Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCysArgIleThr 184
 Db 832 ATCAGCTATTACTTTCCGGCAGTGTGTCAGTGTGGGTCTGTAATTGTCTCGTCTCGTC 891
 Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuLeuThrCysMetGlyIle 204
 Db 892 ACTGCAGCATTTTACTGTAAACATGACGCTCTATCTGTCTCATGACAGTCATAAGCAIT 951
 Qy 205 AsnArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLysLeuProLys 221
 Db 952. GACCGGTTTCTGGTGTGTATCCATGCAGTCCTCTCTCTGGGCTACTCTGGGAAGG 1011
 Qy 222 ArgSerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeu 241
 Db 1012 GCTTCTCTCACT-----TGCTGGCCATCTGGGCTTTGGCCATCGCAGGGGTAGTG 1062
 Qy 242 ProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHis 261
 Db 1063 CCTCTCGTCTCAAGGAGCAAAACCATCCAGGTGCCCGGCTCAACATCACTACTCTGTCA 1122
 Qy 262 AspValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeu 281
 Db 1123 GATGTCTCAATGAACCCCTGCTCGAAGGCTACTATGCTCTACTCTCTCAGCCTCTCT 1182
 Qy 282 AlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeuIle 301
 Db 1183 GCTGTCTTC---TTTTTGTGGCGTGTATTTCCACGGTCTGTATGTGTATGTATCATATT 1239
 Qy 302 HisLysLeuLysSer-----LysAspArgIleTyrLeuGly 313
 Db 1240 CGATGCTTAGTCTTCCGAGTTCACACCGCAGCAAGGTCCTCGGCTTTG----- 1293
 Qy 314 TyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsn 333
 Db 1294 TTCTGTCTGCTGCT-----GTTTCTGCATCTTTCATCATTTGCTTCGGACCCCAAC 1347
 Qy 334 IleIleLeuValIleHisAlaAsnTyrTyrHis-----AsnThrAspSerLeu 351
 Db 1348 GTCTCTGATTCGTCATTAC---TCATTCCTTCTCACACTTCCACACAGAGGTGCC 1404
 Qy 352 TyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 371
 Db 1405 TACTTTCCTTACTCTCTGTGTGTGTGTCAGCAGCATAGTCTGTGTCATCGACCCCTTA 1464
 Qy 372 LeuTyrPheValMetSer 377
 Db 1465 ATTACTATTACGCTTCC 1482

AAZ32191 standard; cDNA; 3472 BP.

AAZ32191;

13-JAN-2000 (first entry)

Human thrombin receptor nucleotide sequence.

Human; coding sequence polymorphism; vascular pathology gene; polymorphic site; phenotype correlation; forensic; paternity testing; medicine; genetic analysis; vascular disease; ds.

Homo sapiens.

WO950454-A2.

07-OCT-1999.

26-MAR-1999; 99WO-US06473.

01-APR-1998; 98US-0054272.

(WHED) WHITEHEAD INST BIOMEDICAL RES.

Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG; WPI; 1999-620066/53.

P-PSDB; AAV49570.

Determination of polymorphisms in genes, especially those identifying predisposition to vascular disease

Claim 1; Fig 35; 134pp; English.

AAZ32159 to AAZ32194 represent reference alleles for specifically claimed nucleic acid sequences from the present invention which comprise polymorphic sites as given in a table in the specification, selected from 92 single nucleotide polymorphisms in which the nucleotide at the polymorphic site is different from a nucleotide at the same site in a reference allele. The nucleic acids, and primers and probes, are used to identify polymorphisms, which may predispose an individual to disease, especially a vascular disease. They can also be used in phenotype correlations, forensics, paternity testing, medicine or genetic analysis. AAY49550 to AAY49573 represent the proteins which correspond to some of the reference alleles.

Sequence 3472 BP; 933 A; 817 C; 785 G; 937 T; 0 other;

Alignment Scores:

Pred. No.:	2.3e-37	Length:	3472
Score:	509.00	Matches:	120
Percent Similarity:	52.73%	Conservative:	83
Best Local Similarity:	31.17%	Mismatches:	140
Query Match:	23.83%	Indels:	42
DB:	20	Gaps:	14

US-09-208-629F-3 (1-407) x AAZ32191 (1-3472)

20 LeulleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle 39

240 CTGCTGTCGTGGTGGCCGCTGCTTC-----AGTCTGTGCGGCGCGCTGTG 284

40 AnValSerAspAsnSerAlaLysPro-----ThrLeuThrIle 52

285 TGTGCCCGCACCCGGGCGCGAGGCCAGATCAAAAGCAACAATGCCACCTTAGATGCC 344

53 LysSerPhe---AsnGlyGlyProGlnAsnThrPheGluPheProLeuSerAspIle 71

345 CGGTCACTTCTCTCAGGAACCCCAATGATAAATGAACCATTTTCGGAGGATGAGGAG 404

72 GluGlyTrpThrGlyAlaThr-----ThrThrIleLysAlaGluCysProGlu 97

405 AAAAATGAAGTGGGTAACTCAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTT 454

Qy	88	AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu	107
Db	465	CAAAAAACAACCTCCTGCATTCATCTCAAGAGATGCCTCCGGATATTGGACACGCTCTCG	524
Qy	108	SerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsn	127
Db	525	CTGACACTCTTTGTCCTCACTGTGTACACGGAGTGTGTGTAGTCAGCTCCCACTAAAC	584
Qy	128	Ile-----ValThrLeuTyrPlysLeuSerLeuArgThrIlyssSerIleSerLeuValIle	145
Db	585	ATCATGCCATCTGTGTGTTCATCTCGAAATGAAGTCAAGAAG---CCGGCGGTGGTG	641
Qy	146	PheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLysIle	165
Db	642	TACATGCTGCCTCGCCACGGCAGATGTCTGTTGTGTCTGTCTCCCTTTAAGATC	701
Qy	166	AlaTyrHisLeuAsnGlyAsnAsnTyrPheGlyGluValMetCysArgIleThrThr	185
Db	702	AGCTATTACTTTCCGGCAGTATTGGCAGTTTGGGTCTGAATGTGTCTCGTCACT	761
Qy	186	ValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsn	205
Db	762	GCAGCATTTTACTGTAAATGTAACGCTCTATCTGTCTATGATGACATCAAGCATTGAC	821
Qy	206	ArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnIlyssLeuProLysArg	222
Db	822	CGGTTTCTGCTGTGTGTATCCCATCCAGTCCCTCTCTCGGGCTACTCTGGGAAGGGCT	881
Qy	223	SerPheSerLeuLeuMetCysGlyIleValTyrValMetValPheLeuTyrMetLeuPro	242
Db	882	TCCTTCACT-----TGCTGGCCATCTGGGCTTTGGCCATCGCAGGGGTAGTCCT	932
Qy	243	PheValIleLeuLysGlnIlyThrHisLeuValHisSerGluIleThrThrCysHisAsp	262
Db	933	CTCGTCTCAAGGAGCAAAACCATCCAGTCCCGGGCTCAACATCACTACCTGTCTCATGAT	992
Qy	263	ValValAspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeuAla	282
Db	993	GTGCTCAATGAACCCCTGCTCGAAGGTACTATGCCCTACTACTTCTCAGGCTTCTCTGCT	1052
Qy	283	PhePheGlyPheLeuIleProPheValIleIlePheCysTyrThrThrLeuIleHis	302
Db	1053	GTCTTC---TTTTTTGGCGCTGATCATTTCCACGGTCTGTTATGTGTCTATCATTTGCA	1109
Qy	303	LysLeuLysSer-----LysAspArgIleTyrPheLeuGlyTyr	314
Db	1110	TGCTTTAGCTCTTTCCGAGTTGCGCAACCGCAGCAAGAAGTCCCGGGCTTTG-----TTC	1163
Qy	315	IleIysAlaValLeuLeuIleValIlePheThrIleCysPheAlaProThrAsnIle	334
Db	1164	CTGTGAGTCTCT-----GTTTCTGCATCTTCATCATTTTGTCTCGACCCCAACAGCT	1217
Qy	335	IleLeuValIleHisHisAlaAsnTyrTyrHis-----AsnThrAspSerLeuTyr	352
Db	1218	CTCCTGATTGGGATTAC---TCATTCCTTTCTCACACTTCCACACAGAGGCTGCCTAC	1274
Qy	353	PheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu	372
Db	1275	TTTGGCTACCTCTGT	1334
Qy	373	TyrPheValMetSer	377
Db	1335	TACTATTACGCTTCC	1349
RESULT 23			
AAZ50771			
ID	AAZ50771	standard; DNA; 3480 BP.	
AC	AAZ50771;		
XX			
DT	31-MAY-2000	(first entry)	
XX			

DT 14-MAR-2001 (first entry)
 XX Human low adenosine antisense oligonucleotide related sequence #2998.
 DE
 XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; bronchodilator; antiinflammatory;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200062736-A2.
 XX
 XX 26-OCT-2000.
 XX
 XX 24-MAR-2000; 2000WO-US08020.
 XX
 XX 06-APR-1999; 99US-0127958.
 XX (UYEC-) UNIV EAST CAROLINA.
 XX (NYCE/) NYCE J W.
 XX
 XX Nyce JW;
 XX
 XX WPI; 2000-679539/66.
 XX
 XX Low adenosine (A) content antisense oligonucleotides which do not
 PT trigger adenosine receptors during metabolism, useful e.g. for treating
 PT cancers and respiratory obstructions -
 XX
 XX Disclosure; Page 1415; 1592pp; English.
 XX
 CC The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
 CC and/or surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 1534 BP; 235 A; 526 C; 479 G; 293 T; 1 other;
 Alignment Scores:
 Pred. No.: 1.09e-37 Length: 1534
 Score: 507.50 Matches: 111

Percent Similarity: 52.97% Conservative: 76
 Best Local Similarity: 31.44% Mismatches: 151
 Query Match: 23.76% Indels: 15
 DB: 21 Gaps: 6
 US-09-208-629f-3 (1-407) x AAF21431 (1-1534)
 QY 27 LeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSerAla 46
 DB 15 CTGCTCTCTGCGCCCTGTGTCTG-----GGTTTCAGCTGTCTGGCGCACCCAG 65
 QY 47 LysProThrLeuThrIleLys-----SerPheAsnGlyGlyProGlnAsnThrPheGlu 64
 DB 66 ACCCCAGCGTCTACGACGAGAGCGGAGCACCGAGGTGGTGTATGATGACGACGCGCTCA 125
 QY 65 GluPheProLeuSerAspIleGluGlyThrPheGlyAlaThrThrThrIleLysAlaGlu 84
 DB 126 ATCTGCTGCTGCCCC-----CGCGCTACCCAGGCCAAGTC----- 161
 QY 85 CysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArg 104
 DB 162 TGTCCCAATGACAGT---GACACCTGGAGCTCCCGGACAGCTCAGCGGCACCTGCTTCTG 218
 QY 105 SerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyVal 124
 DB 219 GGCTGGGTGCCACACGAGGTGTGCCGCCCTCTATGGCTGTGTCTGTGTGGGGCTG 278
 QY 125 ProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuVal 144
 DB 279 CCGGCCAATGGGTGGCGCTGTGGGTGGTGGCCACGACGAGGCACTCGGCTGCCCTCCACC 338
 QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
 DB 339 ATGCTCTGATGACCTCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 398
 QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
 DB 399 ATCGCTACACCTGCTGGTGGCGAGCGCTGCTGGCGAGGCGCGCTGCGCTGCGCTGCGCC 458
 QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleThrCysMetGlyIle 204
 DB 459 ACGGCGCGACTCTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 518
 QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
 DB 519 GATCGTACTGGCCCTGTGTGACCGCTGCGGGCGCGCGCTGCGTGGCGCGCGCTG 578
 QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
 DB 579 GCCCTTGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638
 QY 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
 DB 639 CTACAGCGGAGACCTTCGGGTGGCGCTGCGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698
 QY 265 AspAlaCysGluSerProSerPheArgPheTyrTyrPheValSerLeuAlaPhePhe 284
 DB 699 ---CCCTTGGACGACAGGCTCCCACTGGCAACCGGCTTCACTGCTGCTGCTGCTGCTGCTG 755
 QY 285 GlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrThrLeuIleHisLysLeu 304
 DB 756 GGCTGTCTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 815
 QY 305 LysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValIleLeuLeuValIle 324
 DB 816 GCGGCGAGCGCGCGCTACCGGCCACGCGCTGAGCTGACCGAGTGGTGGTGGTGGTGGTGGT 875
 QY 325 PheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyr 344
 DB 876 GCGGTGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 935
 QY 345 TyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeu 364

Db 936 CCCAGCCCTGGGGCAACCTCTATGGTGCTACGTGCCCGCCAGCCTGGCGCTGAGCACCCCTC 995

Qy 365 AsnSerCysLeuAspProPheLeuTyrPheValMetSer 377

Db 996 AACAGCTGCTGGATGCCCTTCATCTACTACTACGTGTCG 1034

RESULT 26

AAA35309

ID AAA35309 standard; DNA; 1534 BP.

XX AC AAA35309;

XX 28-JUL-2000 (first entry)

XX Human adenosine receptor related polynucleotide 2nd SEQ ID NO:183.

XX Human; adenosine receptor; low adenosine antisense oligonucleotide;

XX phosphorothioate; impaired respiration; inflammation; allergy;

XX allergic disease; bronchoconstriction; inhibitor; antiinflammatory;

XX antiallergic; antiasthmatic; cytotatic; analgesic; impaired airway;

XX lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;

XX respiratory distress syndrome; pain; cystic fibrosis; emphysema;

XX pulmonary hypertension; chronic obstructive pulmonary disease; COPD;

XX cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX Homo sapiens.

XX WO200009525-A2.

XX 24-FEB-2000.

XX 03-AUG-1999; 99WO-US1712.

XX 03-AUG-1998; 98US-0095212.

XX (UYEC-) UNIV EAST CAROLINA.

XX Nyce JW;

XX WPI; 2000-205971/18.

XX New antisense oligonucleotides useful for treating e.g. pulmonary

XX vasoconstriction, inflammation, allergies, asthma, hypertension,

XX bronchitis, emphysema, respiratory distress syndrome, ischemia or

XX cancers

XX Disclosure; Page 1330; 1343pp; English.

XX The present invention describes a new composition comprising an

XX antisense oligonucleotide (ON) with low adenosine (up to 15%), which

XX targets nucleic acids involved in bronchoconstriction, allergies, and/or

XX inflammation. The ON can have antiinflammatory, antiallergic,

XX antiasthmatic, cytotatic and analgesic activities. The compositions are

XX useful for the treatment of diseases associated with inflammation,

XX impaired airways, including lung disease and diseases whose secondary

XX effects afflict the lungs of a subject. They can be used for treating

XX e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,

XX asthma, impeded respiration, respiratory distress syndrome, pain, cystic

XX fibrosis, pulmonary hypertension, emphysema, chronic obstructive

XX pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,

XX carcinomas, and cancers which may metastasise to the lungs, including

XX breast and prostate cancer. The reduction of the adenosine content of

XX the ONs reduces side effects. The A-containing ONs break down with the

XX release of deoxyadenosine which activates adenosine receptors causing

XX bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the

XX nucleotide sequences given in the sequence listing from the present

XX invention, which correspond to SEQ ID NO:1 to 2815, and then the last

XX 185 sequences are also called SEQ ID NO:1 to 185, but the sequences

XX differ from the previously named sequences. SEQ ID NO:11 to 1680

XX (AAA32323 to AAA33992) are specifically claimed. SEQ ID NO:1 from the present

XX invention. N.B. Sequences given in the disclosure of the present

XX invention do not match up with their corresponding SEQ ID NO: sequences

XX given in the sequence listing.

XX SQ Sequence 1534 BP; 235 A; 526 C; 479 G; 293 T; 1 other;

Alignment Scores:

Prod. No.: 1.09e-37 Length: 1534

Score: 507.50 Matches: 111

Percent Similarity: 52.97% Conservative: 76

Best Local Similarity: 31.44% Mismatches: 151

Query Match: 23.76% Indels: 15

DB: 21 Gaps: 6

US-09-208-629F-3 (1-407) x AAA35309 (1-1534)

Qy 27 LeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSerAla 46

Db 15 CTGCTCTCTGGCCCTGTGTGCTG-----GGTTTCAGCTGTCTGGCGCACCCAG 65

Qy 47 LysProThrLeuThrIleLys-----SerPheAsnGlyGlyProGlnAsnThrPheGlu 64

Db 66 ACCCCAGCGCTTACGACGAGCGGAGCACCGGAGGTGGTGATGACAGCACCGCCCTCA 125

Qy 65 GluPheProLeuSerAspIleGluGlyTyrThrGlyAlaThrThrThrIleLysAlaGlu 84

Db 126 ATCTGCTGCTGCCCC-----CGGGCTACCCAGGCCAAGTC----- 161

Qy 85 CysProGluAspSerIleSerThrIleuHisValAsnAsnAlaThrIleGlyTyrLeuArg 104

Db 162 TGTGCCAATGACAGT--GACACCTGGAGCTCCCGACAGCTCACGGGCACCTGCTTCG 218

Qy 105 SerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyVal 124

Db 219 GGTGGTGGTCCCAACAGGCTGGTGGCCGCTTATGGGCTGGTCTGGTGGTGGGCTG 278

Qy 125 ProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuVal 144

Db 279 CCGGCCAATGGCTGGCGCTGGGTGGTGGCCAGCGAGCCTCGGCTGGCTCCCTCCACC 338

Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164

Db 339 ATGTGCTGATGAACCTCGCGACTGTGACCTCTGTGCGCTGGCGCTGCCCCCGCG 398

Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTyrPheGlyGluValMetCysArgIleThr 184

Db 399 ATCGCTACACCTGCTGGTGGCGAGCGCTGGCGCTTCGGGGAGCGCCCTCGCGCTGCC 458

Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204

Db 459 ACGGCGCACTCTATGGTCACATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 518

Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224

Db 519 GATCGCTACTGGCCCTGTGTGCGCGCTGGCGCGCGCGCTGGCTGGCGCGCGCTG 578

Qy 225 SerLeuLeuMetCysGlyIleValTyrPheValMetValPheLeuTyrMetLeuProPheVal 244

Db 579 GCGCTTGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638

Qy 245 IleLeuGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264

Db 639 CTGACGGCGCACCTTCCGGCTGGCGCGCTCCGATCGGCTGCTGCTGCTGCTGCTGCTGCTG 698

Qy 265 AspAlaCysGluProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePhe 284

Db 699 ---CCCTGAGCGACAGCCCTCCACCTGGCAACCGCCCTTCACCTGCTGCTGCTGCTGCTG 755

Qy 285 GlyPheLeuIleProPheValIleIlePheCysTyrThrThrThrLeuIleHisLysLeu 304

Db 756 GGTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 815

Qy 305 LysSerLysAspArgIleThrPheGlyTyrIleLysAlaValLeuLeuLeuValIle 324

Db 816 GCGGCGAGCGCGCGCTGACGCGCTGAGGCTGACCGCTGAGGCTGAGGCTGAGGCTGAGGCTG 875

QY 325 PheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyr 344
 Db 876 GCCGTGGCCTTCTCGCCAGCAACTGCTGCTGCTGCTACTCGGACCCGAGC 935
 QY 345 TyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeu 364
 Db 936 CCCAGGCCCTGGGGCAACCTCTATGGTGCCTACGTGCGCCAGCCTGGCGCTGAGCACCCTC 995
 QY 365 AsnSerCysLeuAspProPheLeuTyrPheValMetSer 377
 Db 996 AACAGCTGCGTGGATCCCTTCATCTACTACTACGTGTCG 1034
 RESULT 27
 AA232747
 ID AA232747 standard; cDNA; 4895 BP.
 XX AC AA232747;
 XX 31-JAN-2000 (first entry)
 XX Human protease-activated receptor. PAR4 cDNA.
 XX
 KW Protease-activated receptor; PAR4; ZCHEMR2; G protein coupled;
 KW cellular signalling; protease; tethered ligand; N-terminal;
 KW proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist;
 KW antagonist; cellular response; physiological response; clotting pathway;
 KW platelet; proliferation; differentiation; mediation;
 KW inflammatory process; vascular injury; chemotaxis; mitogenesis;
 KW growth factor; production; ds.
 XX Homo sapiens.
 OS
 PH Location/Qualifiers
 FT CDS
 FT 176..1333
 FT /*tag= a
 FT /product= "Human protease-activated receptor PAR4"
 FT 5'UTR
 FT 1..175
 FT /*tag= b
 FT 176..226
 FT /*tag= c
 FT 227..1330
 FT /*tag= d
 FT /product= "Mature human PAR4"
 FT 3'UTR
 FT 1333..4895
 FT /*tag= e
 FT /note= "GC-rich"
 XX WO9950415-A2.
 XX 07-OCT-1999.
 XX 31-MAR-1999; 99WO-US07100.
 XX 01-APR-1998; 98US-0053866.
 XX (Zymo) ZYMOGENETICS INC.
 XX (UNIW) UNIV WASHINGTON.
 XX Xu W, Presnell SR, Yee DP, Foster DC;
 XX WPI; 1999-633640/54.
 XX P-PSDB; RAY50135.
 XX Novel protease activated receptor 4, useful for screening for
 PT (ant)agonists for promoting the proliferation and/or differentiation of
 PT platelets and in mediating inflammatory events
 XX Claim 1; Page 76-80; 85pp; English.
 XX This sequence represents a human protease-activated receptor PAR4 (also
 CC referred to as ZCHEMR2) cDNA which was identified from EST (expressed
 CC sequence tag) sequences with homology to the three known protease-
 CC activated receptors (PAR1, PAR2 and PAR3). Protease-activated receptors

CC (PARs) are a subfamily of G protein coupled receptors which are capable
 CC of mediating cellular signalling in response to proteases (e.g.,
 CC thrombin). They are characterised by a tethered peptide ligand at the
 CC extracellular N-terminus that is generated by proteolysis. PAR4 is
 CC activated by thrombin or trypsin cleavage at Arg47/Gly48, which generates
 CC a new N-terminus corresponding to the tethered ligand (a hexapeptide).
 CC Agonists of PAR4 are useful for upregulating cellular or physiological
 CC responses whereas antagonists are used to downregulate these
 CC activities. The PAR4 protein is further useful for dissecting the
 CC effects of thrombin or other activating proteases in the clotting
 CC pathway from the effects of these proteases at the cellular level.
 CC Agonists are specifically useful in promoting the proliferation
 CC and/or differentiation of platelets, in mediating inflammatory events,
 CC responses to vascular injury, chemotaxis or mitogenesis, and in
 CC producing growth factors. Antagonists are useful as research reagents
 CC for characterising sites of ligand-receptor interaction.
 XX
 SQ Sequence 4895 BP; 971 A; 1573 C; 1400 G; 951 T; 0 other;
 Alignment Scores:
 Pred. No.: 4,97e-37 Length: 4895
 Score: 507.50 Matches: 111
 Percent Similarity: 52.97% Conservative: 76
 Best Local Similarity: 21.44% Mismatches: 151
 Query Match: 23.76% Indels: 15
 DB: 20 Gaps: 6
 US-09-208-629F-3 (1-407) x AA232747 (1-4895)
 QY 27 LeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSerAla 46
 Db 188 CTGCTCTCTGGGCCCTCTGGTCTG-----GGGTTCAGCCTGCTGCGGCACCCAG 238
 QY 47 LysProThrLeuThrIleLys-----SerPheAsnGlyGlyProGlnAsnThrPheGlu 64
 Db 239 ACCCCAGGCTCTACGACGAGCGGAGTGGTGTGATGACGACGACGCGCTCA 298
 QY 65 GluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrThrIleLysAlaGlu 84
 Db 299 ATCTCTGCTGCCCTCC-----CGGGCTACCCAGGCCAAGTC----- 334
 QY 85 CysProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArg 104
 Db 335 TGTGCCAATGACAGT--GACACCTGGAGCTCCGCGACAGCTCACGGGCACTGCTTCG 391
 QY 105 SerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyVal 124
 Db 392 GGCTGGGTGCCACACGAGCTGGTCCGCCCTCTATGGGTGGTCTCTGGTGGGGCTG 451
 QY 125 ProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuVal 144
 Db 452 CCGGCCAATGGGCTGGGCTGTGGTGTGGTGTGGCCAGCAGCAGCTGGCTGCCCTCCACC 511
 QY 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
 Db 512 ATGCTGCTGATGAACCTCGGACTGCTGACCTCTCTGTGGCCCTGGCGCTGCCCGCGG 571
 QY 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
 Db 572 ATCGCTTACCACCTGGTGGCCAGCGCTGCGCCCTTCGGGGAGCGCGCTGCCCTCGGCC 631
 QY 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
 Db 632 ACCGGCGCACTCATGTGTACATGTATGGCTAGTGTGTGTGGCCCGCGCTGAGCTG 691
 QY 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
 Db 692 GATCGCTACCTGGCCCTGGTGGACCCGCTGCGGGCGCGCCCTGCGTGGCGCGCGCTG 751
 QY 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuPheVal 244
 Db 752 GCCCTTGGACTCTGATGGTGTGTGTGCTCATGGCGCGCGCCCTGGCACTGCCCTGACA 811

Db 793 GCCTTCTCTGCTGCTTC---TTTTTGTGCGCTGATCATTTCCACGGTCTGTATGTG 849
 Qy 299 ThrLeuIleHisIysLeuLysSer-----LysAspArgIle 310
 Db 850 TCTATCATTCGATGCTTGTAGCTCTTCGCGAGTTGCCAACCGCAGCAAGAGTCCCCGGCT 909
 Qy 311 TrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrIleCysPheAla 330
 Db 910 TTG-----TTCGTCTCAGTGTCT-----GTTTCTGCACTCTCATCTTTCCTTCGGA 957
 Qy 331 ProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyrHis-----AsnThr 348
 Db 958 CCCACAAACGTCCTCCTGATTGGCAATTAC---TCATTCTCTTCTCACACTTCCACCACA 1014
 Qy 349 AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
 Db 1015 GAGCTCCCTACTTTGCTCTACCTCCTCTGCTGTCTGTCTGTCAGCAGCATAGCTCGTGATC 1074
 Qy 369 AspProPheLeuTyrPheValMetSer 377
 Db 1075 GACCCCTAATTACTATTACGCTTCC 1101

RESULT 31

AAF21432
 ID AAF21432 standard; DNA; 3182 BP.

AC AAF21432;
 XX
 XX
 DT 14-MAR-2001 (first entry)
 XX
 XX
 DE Human low adenosine antisense oligonucleotide related sequence #2999.
 XX
 XX
 KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; ROS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200062736-A2.
 XX
 PD 26-OCT-2000.
 XX
 PF 24-MAR-2000; 2000WO-US08020.
 XX
 PR 06-APR-1999; 99US-0127958.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 PA (NYCE/) NYCE J W.
 XX
 PI Nyce JW;
 XX
 DR WPI; 2000-679539/66.
 XX
 PT Low adenosine (A) content antisense oligonucleotides which do not
 trigger adenosine receptors during metabolism, useful e.g. for treating
 cancers and respiratory obstructions -
 PT
 XX
 PS Disclosure; Page 1415-1416; 1592pp; English.
 XX
 CC The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the

CC expression and or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adenosine molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
 CC and/or surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.
 XX
 XX

SQ Sequence 3182 BP; 895 A; 701 C; 684 G; 902 T; 0 other;

Alignment Scores:
 Pred. No.: 9,21e-37 Length: 3182
 Score: 502.00 Matches: 113
 Percent Similarity: 54.15% Conservative: 76
 Best Local Similarity: 32.38% Mismatches: 130
 Query Match: 23.50% Indels: 30
 DB: 21 Gaps: 12

US-09-208-629F-3 (1-407) x AAF21432 (1-3182)

Qy 49 ThrLeuThrIleLysSerPhe---AanGlyGlyProGlnAsnThrPheGluGluPhePro 67
 Db 43 ACCTTAGATCCCGGTCAITTTCTCTCAGAAACCCCAATGATAAATATGAACCATTTGG 102
 Qy 68 LeuSerAspIleGluGlyTrpThrGlyAlaThr-----ThrThrIleLysAla 83
 Db 103 GAGGATGAGGAGAGAAATAAGTGGTTAACTGATACAGATTAGTCTCCATCAATAA 162
 Qy 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
 Db 163 AGCAGTCTCTTCAAAAACAACTTCTCGCATTCATCTCAGAGATCCCTCCGGAATTTG 222
 Qy 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGly 123
 Db 223 ACCAGTCTCTGGCTGACACTTTTGTCCCATCTGTGTACACCGGAGTGTGTGTAGTCAGC 282
 Qy 124 ValProSerAsnIle-----ValThrLeuTrpLysLeuSerLeuArgThrLysSerIle 141
 Db 283 CTCCCACTAAACATCATGCGCATCGTTGTGTTCATCTGAAATGAAGGTCAAGAAG--- 339
 Qy 142 SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeu 161
 Db 340 CCGCGGTGTGTGTACATGTGCACCTGGCCACGCGCAGATGTGTGTGTGTGTGTGTGCTC 399
 Qy 162 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCys 181
 Db 400 CCCTTTAAGATCAGCTATTACTTTTCCGGCAGTGATGGCAGTTGGGTCTGAATTTGT 459
 Qy 182 ArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCys 201
 Db 460 CGTTCGTCTCAGTCAGCATTTTACTGTAAACATGTACGCCCTCTATCTTGTCTCATGACAGC 519
 Qy 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisPro-----PheThrTyrGlnLys 218
 Db 520 ATAAGCATTTGACCGGTTTCTGGCTGTGGTGTATTCATCCATGACGATCCCTCTCTGGGCTACT 579
 Qy 219 LeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeu 238

```
Db 580 CTGGAGAGGCTCTCTTCACT-----TGCTGCCCATCTGGGCTTTGGCGCATGCGA 630
Qy 239 TyrMetLeuProPheValIleLeuLeuGlnGluTyrHisLeuValHisSerGluIleThr 258
Db 631 GGGGTAGTGCCTCTCGTCTCAAGGAGCAACCATCCAGGTGCCGGGCTCAACATCACT 690
Qy 259 ThrCysHisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPhe 278
Db 691 ACCTGTATGATGTCTCAATGAACCCCTGCTCGAAGCTACTATGCTACTACTACTCTCA 750
Qy 279 ValSerLeuAlaPhePheGlyPheLeuLeuProPheValIleIleIlePheCysTyrThr 298
Db 751 GCCTTCTCTGCTCTTC-----TTTTTGTGCCGTGATCAATTCACGGTCTGTATGTG 807
Qy 299 ThrLeuIleHisLeuLeuSer-----LysAspArgIle 310
Db 808 TCTATCATTCGATGCTCTTACTCTTCGCGAGTTGCCAACCGCAGCAAGAGTCCGGGCT 867
Qy 311 TrpLeuGlyTyrIleLeuAlaValLeuLeuLeuValIlePheThrIleCysPheAla 330
Db 868 TTG-----TTCCTGTGAGTGTCT-----GTTTCTGCATCTTCATCATTTGCTTGGGA 915
Qy 331 ProThrAsnIleIleValIleHisHisAlaAsnTyrTyrTyrHis-----AsnThr 348
Db 916 CCACAAACGTCCTCTGATGCGCAATAC---TCATTCCTTTCTCAGCTTCCACACACA 972
Qy 349 AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 973 GAGGCTGCCTACTTGTGCTACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1032
Qy 369 AspProPheLeuTyrPheValMetSer 377
Db 1033 GAGCCCCCTAATTACTATTACGCTTC 1059
RESULT 32
AA35310
ID AAA35310 standard; DNA; 3182 BP.
XX
AC AAA35310;
XX
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:184.
XX
KW Human; adenosine receptor; low adenosine antiseize oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antisthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
XX WO200009525-A2.
XX
XX 24-FEB-2000.
XX
XX 03-AUG-1999; 99WO-US17712.
XX
XX 03-AUG-1998; 98US-0095212.
XX
XX (UYEC-) UNIV EAST CAROLINA.
XX
XX Nyce JW;
XX
XX WPI; 2000-205971/18.
XX
XX New antiseize oligonucleotides useful for treating e.g. pulmonary
XX vasoconstriction, inflammation, allergies, asthma, hypertension,
XX bronchitis, emphysema, respiratory distress syndrome, ischemia or
```

cancers -

Disclosure; Page 1330-1331; 1343pp; English.

The present invention describes a new composition comprising an antiseize oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antisthmatic, cytostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasize to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing.

SQ Sequence 3182 BP; 895 A; 701 C; 684 G; 902 T; 0 other;

Alignment Scores:

Pred. No.:	9-21e-37	Length:	3182
Score:	502.00	Matches:	113
Percent Similarity:	54.15%	Conservative:	76
Best Local Similarity:	32.38%	Mismatches:	130
Query Match:	23.50%	Indels:	30
DB:	21	Gaps:	12

US-09-208-629F-3 (1-407) x AAA35310 (1-3182)

```
Qy 49 ThrLeuThrIleLysSerPhe---AenGlyGlyProGlnAsnThrPheGluGluPhePro 67
Db 43 ACCTTAGATCCCGGTCAITTTCTCTCAGAACCCCAATGATAAATATGAACCAATTTGG 102
Qy 68 LeuSerAspIleGluGlyTrpThrGlyAlaThr-----ThrThrIleLysAla 83
Db 103 GAGGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 162
Qy 84 GluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeu 103
Db 163 AGCAGTCTCTTCAAAAACAACTTCTGTCATCTCAGAGATGCTCCGATATTG 222
Qy 104 ArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValGly 123
Db 223 ACCAGCTCTGGCTGACACTTTTGTCCATCTGTGTACACCGAGTGTGTGTGTGTGTGT 282
Qy 124 ValProSerAsnIle-----ValThrLeuTyrLysLeuSerLeuArgThrLysSerIle 141
Db 283 CTCCCACTAAACATCATGCGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 339
Qy 142 SerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeu 161
Db 340 CCGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 399
Qy 162 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCys 181
Db 400 CCCTTTAAGATCAGCTATTACTTTTCCCGCAGGTGATGGCAGTGTGGTGTGTGTGTGTGT 459
Qy 182 ArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCys 201
Db 460 CGTTCGTCACTGCAGCATTTTACTTAACATGTACGCTCTATCTATCTGTGTGTGTGTGT 519
```


Db 658 TCCTTCACT-----TGCTGGCCATCTGGCTTTGGCCATCGCAGGGGTAGTGCCCT 708
 Qy 243 PheValIleLeuLysGlnGluTyHisLeuValHisSerGluIleThrCysHisAsp 262
 Db 709 CTGCTCCTCAAGGAGCAACCATCCAGGTGCCGGGCTCAACATCACTACTGTCATGAT 768
 Qy 263 ValValAspLacYsGluSerProSerPheArgPheTyTyPheValSerLeuAla 282
 Db 769 GTGCTCAATGAACCCCTGCTCGAAGGCTACTATGCTCTACTACTTCTCAGCCTTCTCTGCT 828
 Qy 283 PhePheGlyPheLeuIlePheProPheValIleIlePheCysTyThrThrLeuLeuHis 302
 Db 829 GTCTTC---TTTTTGTGGCGCTGATCATTTCCACGGCTGTGTATGTCTATCATTCGA 885
 Qy 303 LysLeuLysSer-----LysAspArgIleTyrLeuGlyTyr 314
 Db 886 TGTCTTAGCTCTTCGGCAGTTGCCAACCGCAGCAAGAAGTCCCGGGCTTTG-----TTC 939
 Qy 315 IleLysAlaValLeuLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 940 CTGTCACTGCT-----GTTTCTGCTATCTTTCATCATTTGCTTCGGACCCCAACGTC 993
 Qy 335 IleLeuValIleHisHisAlaAsnTyrTyTyrHis-----AsnThrAspSerLeuTyr 352
 Db 994 CTCTGATGGCATTCAC---TCATCTTCTTTCACACTTCCACCACAGAGGTGCTAC 1050
 Qy 353 PheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu 372
 Db 1051 TTTGCTTACCTCTCTGTGTGTGTGTCAGCAGATAGCTGCTGCATCGACCCCTAATT 1110
 Qy 373 TyrPheValMetSer 377
 Db 1111 TACTATTACGCTTCC 1125

RESULT 34
 AAH51006
 ID AAH51006 standard; DNA; 2137 BP.
 XX
 AC AAH51006;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE Human nGPCR5 coding sequence #2.
 XX
 KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disorder;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder/attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective; ds.
 XX
 OS Homo sapiens.
 XX
 PN WC200136473-A2.
 XX
 XX 25-MAY-2001.
 XX
 PF 16-NOV-2000; 2000WO-US31581.
 XX
 PR 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166678.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0186530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.

PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX (PHAA) PHARMACIA & UPJOHN CO.
 PA
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX
 XX WPI; 2001-389826/41.
 DR P-PSDB; AAG80966.
 DR
 XX
 PT New G protein-coupled receptor (nGPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 XX
 XX Claim 4; Page 88; 261pp; English.
 PS
 XX
 CC The present invention relates to novel G protein-coupled receptors
 CC (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
 CC sequence is the coding sequence for one such G protein-coupled receptor.
 CC GPCRs are also known as seven transmembrane receptors and function in
 CC signal transduction. The nGPCRx coding sequences are useful for
 CC screening a human to diagnose a disorder affecting the brain or a genetic
 CC predisposition, specifically schizophrenia. nGPCRx are useful for
 CC identifying compounds useful for treating schizophrenia. Detection of
 CC nGPCRx in a sample is useful as a diagnostic tool for diseases or
 CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,
 CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular
 CC diseases, proliferative disorders and hormonal disorders. Modulators of
 CC nGPCRx activity have the utility for treating neurological disorders,
 CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity
 CC disorder/attention deficit disorder), and neuronal disorders such as
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.
 CC Additional disorders include inflammatory conditions (e.g. Crohn's
 CC disease), rheumatoid arthritis, autoimmune disorders, cancers,
 CC respiratory ailments such as asthma, and inflammatory diseases e.g.
 CC inflammatory bowel disease.
 XX
 SQ Sequence 2137 BP; 344 A; 719 C; 607 G; 467 T; 0 other;

Alignment Scores:

Pred. No.: 8,4e-37 Length: 2137
 Score: 500.00 Matches: 113
 Percent Similarity: 52.12% Conservative: 59
 Best Local Similarity: 34.24% Mismatches: 128
 Query Match: 23.41% Indels: 30
 DB: 22 Gaps: 8

US-09-208-629f-3 (1-407) x AAH51006 (1-2137)

Qy 69 SerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAsp 88
 Db 150 TCTGAGTCTCACTGCTACTTCTGCGGTCTCT-----GCACAGACCCGGCGAG 203
 Qy 89 SerIleSerThrLeuHisVal-----AsnAsnAlaThrIleGlyTyr 102
 Db 204 GACCCCTCCAGGATGAGGTCCGGAACAGACACCGCGCCCGACACGCGCTGCAGATG 263
 Qy 103 LeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuPheValVal 122
 Db 264 CTGCGGAACCCGGCGATCGGGTGGCCCTCCCGCTGGTGTACTCGTGGTGGCGGGTTC 333
 Qy 123 GlyValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrIleSer 142
 Db 324 AGCATCCCGGGCAACCTCTTCTGTGGTCTGTGCGGGCATGGGGCCAGATCC 383
 Qy 143 ---LeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeu 161
 Db 384 CCGTCGGTCTATCTTCATGATCAACCTGAGCGTCACGACCTGATGTGGCCAGCGGTGTTG 443


```

Db 576 AACGTGGTACCGTGGCGCTTTTACGCAACATGTATTCCAGCATCTCCACCATCACTGTT 635
Qy 202 MetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLys 221
Db 636 ATCAGCGTGGAGCGCTTCTGGGGTCTGTACCGCTCAGTCCAAGCGTGGCGCGC 695
Qy 222 ArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeu 241
Db 696 CGTGGTTACCGGGTGGCGCGTGTGTGAGGACCTGGCTGTCTCTCCTGACCGCGCTGTC 755
Qy 242 ProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHis 261
Db 756 CCGTGGCGCCACCGATCTCACCTACCGGTGACCGCTGGGCATCATCACTGCTTC 815
Qy 262 AspValValAspAlaCysGluSerProSer-----SerPheArgPheTyr 276
Db 816 GACGTCTCAAGTGCAGCATCTCTCCCGAGCGTGGCCATGTGGCGCGTTCCTCTCACC 875
Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCys 296
Db 876 ATCTTCATCTCTGCTG-----TTCCTCATCCCGTTCTGTGATCACCAGTGGCTGT 923
Qy 297 TyrThrThrLeuIleHisLysLeuLysSerLysAspArgIleTrp----- 311
Db 924 TACACGGCCACCATCTCAAGCTTGTGGCAGGAGGAGCGCACGCGCGGAGCAGCGG 983
Qy 312 LeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaPro 331
Db 984 AGCGCCGCGTGGCGCTGGCGCGTGTGTGTGTGCGCTTGTACCTGCTTGCCTCC 1043
Qy 332 ThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeu 351
Db 1044 AACAACTTCGTGCTCTGGCGCACATCGTACGCGCGCTGTCTACGGC---AAGAGCTAC 1100
Qy 352 TyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 371
Db 1101 TACCAGGTGTACAAGCTCAGCTGCTGTCTAGTGCCTCAACAACACTGTCTGGACCGGTT 1160
Qy 372 LeuTyrPheValMetSerLys 378
Db 1161 GTTATTACTTTGGTCCCGG 1181

RESULT 36
AAD01131
ID AAD01131 standard; cDNA; 1080 BP.
XX AC AAD01131;
XX DT 02-NOV-2000 (first entry)
XX DE Human orphan G protein-coupled receptor hCHN4 cDNA.
XX KW Human; orphan G protein-coupled receptor; GPCR; hCHN4; drug screening;
XX KW transmembrane receptor; expressed sequence tag; EST; signal cascade; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 1..1080
XX FT /*tag= a
XX FT /product= "hCHN4"
XX FT /note= "Human orphan G protein-coupled receptor"
XX FN WO200031258-A2.
XX PD 02-JUN-2000.
XX PF 13-OCT-1999; 99WO-US23687.
XX PR 20-NOV-1998; 98US-0109213.
XX PR 16-FEB-1999; 99US-0120416.
XX PR 26-FEB-1999; 99US-0121852.
XX PR 12-MAR-1999; 99US-0123946.

```

```

PR 12-MAR-1999; 99US-0123949.
PR 28-MAY-1999; 99US-0136436.
PR 28-MAY-1999; 99US-0136437.
PR 28-MAY-1999; 99US-0136439.
PR 28-MAY-1999; 99US-0136567.
PR 28-MAY-1999; 99US-0137127.
PR 28-MAY-1999; 99US-0137131.
PR 28-JUN-1999; 99US-0141448.
PR 29-SEP-1999; 99US-0156555.
PR 29-SEP-1999; 99US-0156633.
PR 29-SEP-1999; 99US-0156634.
PR 23-SEP-1999; 99US-0156853.
PR 01-OCT-1999; 99US-0157280.
PR 01-OCT-1999; 99US-0157281.
PR 01-OCT-1999; 99US-0157282.
PR 01-OCT-1999; 99US-0157293.
PR 01-OCT-1999; 99US-0157294.
PR 12-OCT-1999; 99US-0416760.
PR 12-OCT-1999; 99US-0417044.
XX PA (AREN-) ARENA PHARM INC.
XX PI Chen R, Dang HT, Liaw CW, Lin I;
XX DR WPI; 2000-400068/34.
XX DR P-PSDB; AAY71304.
XX PT Novel human orphan G protein-coupled receptors and the encoding cDNAs
XX PS for use in the identification of G protein-coupled receptor agonists -
XX PS Claim 53; Page 76; 102pp; English.
XX CC The present sequence is a cDNA encoding hCHN4, an endogenous human
XX CC orphan G protein-coupled receptor (GPCR). The hCHN4 cDNA was identified
XX CC using full length EST (expressed sequence tag) 1184934 as a probe.
XX CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane
XX CC alpha helices with an extracellular N-terminus and an intracellular
XX CC C-terminus. However, no endogenous ligands has yet been identified for
XX CC the proteins of the invention. The orphan GPCRs may be used in the
XX CC identification of their endogenous ligands, and to screen potential GPCR
XX CC agonists and antagonists for use as pharmaceutical agents. The proteins
XX CC may also be used in the study of GPCR-mediated signalling cascades, and
XX CC to elucidate their precise role in normal and diseased human conditions.
XX CC Nucleic acid encoding human orphan GPCRs may be used for tissue
XX CC localisation expression analysis to provide information about their
XX CC function in healthy and pathological states.
XX SQ Sequence 1080 BP; 162 A; 369 C; 322 G; 227 T; 0 other;

Alignment Scores:
Pred. No.: 1,25e-36 Length: 1080
Score: 494.00 Matches: 106
Percent Similarity: 53.54% Conservative: 53
Best Local Similarity: 35.69% Mismatches: 116
Query Match: 23.13% Indels: 22
DB: 21 Gaps: 6

US-09-208-629F-3 (1-407) x AAD01131 (1-1080)
Qy 96 AsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIle 115
Db 28 GACAACGCGACGCTGCAGATGCTGCGGAACCCCGCGATCGCGTGGCCCTGCCCGTGGTG 87
Qy 116 TyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeuSer 135
Db 88 TACTCGTGTGGCGCGGTGACATCCCGGGCAACCTCTCTCTGTGGGTGCTGTC 147
Qy 136 LeuArgThrLysSerIleSer---LeuValIlePheHisThrAsnLeuAlaIleAlaAsp 154
Db 148 CGCGCGATGGGGCCAGATCCCGTCCGTCATCTTCATGATCACTGACGTCACCGAC 207
Qy 155 LeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 174

```

Db 208 CTGATGCTGGCCAGCGTGTGGCTTTCCAAATCTACTACCATTGCAACCCACCACCTGG 267
Qy 175 ValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCys 194
Db 268 GTATTCGGGTGCTGCTTTGCAACGCTGGTACCGTGGCCCTTTACGCCAACATGTTATCC 327
Qy 195 AlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPhe 214
Db 328 AGCATCTCACCATGACCTGATCAGCGTGGAGCGCTCTCTGGGGGCTCTGTACCCGCTC 387
Qy 215 ThrTyrGlnIlyLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyrVal 234
Db 388 AGTCCACGCGCTGGCCCGCGCTGTTACCGGTGGCCGCTGTCAGGAGCACTGGCTC 447
Qy 235 MetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHis 254
Db 448 CTGCTCTGACCGCCCTGTGGCCCTGGCGGCGCACCGATCTCACCCTACCCGGTGCAGCC 507
Qy 255 SerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSer----- 271
Db 508 CTGGCATCATCATCTGCTTCGACGCTCTCAAGTGGACGATGCTCCCGGCGTGGCCATG 567
Qy 272 -----SerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIlePro 289
Db 568 TGGGCGGTGTTCTCTTCACCATCTTCATCTGCTG-----TTCCTCATCCCG 615
Qy 290 PheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeu----- 304
Db 616 TTCTGATACCGGTGTTTACCGCCACCATCTCAAGCTTGTGGCCAGGAGGAG 675
Qy 305 -----LysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuIle 321
Db 676 GCGCACGCGCGGAGCAGCGGCGGCTGGCGGCTGGCGGCTGCTGCTG--- 732
Qy 322 LeuValIlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAla 341
Db 733 -----GCCTTTGTACCTGCTTCGCCCCAACCACTTCGCTGCTCTGGCGCACATCGT 786
Qy 342 AsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeu 361
Db 787 AGCGCGCTGTCTACGGC---AAGAGCTACTACACGCTGTACAGCTCAGCTGTGTCTC 843
Qy 362 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 844 AGCTGCTCAACAACCTGCTCGACCGCTTGTATTACTTTGCTGCTCCCG 894

RESULT 37

AAA46032
ID AAA46032 standard; cDNA; 1080 BP.

AC AAA46032;

DT 22-AUG-2000 (first entry)

DE Human G protein coupled receptor hCHN4 encoding cDNA SEQ ID NO:29.

XX Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW identification; agonist; screening; therapeutic; pharmaceutical;
KW mutant; SS.

OS Homo sapiens.

XX WO200022131-A2.

PN 20-APR-2000.

XX 13-OCT-1999; 99WO-US24065.

PF 13-OCT-1998; 98US-0170496.

PR 12-NOV-1998; 98US-0108029.

PR 20-NOV-1998; 98US-0109213.

PR 27-NOV-1998; 98US-0110060.

PR 16-FEB-1999; 99US-0120416.

PR 26-FEB-1999; 99US-0121852.
PR 12-MAR-1999; 99US-0123944.
PR 12-MAR-1999; 99US-0123945.
PR 12-MAR-1999; 99US-0123946.
PR 12-MAR-1999; 99US-0123948.
PR 12-MAR-1999; 99US-0123949.
PR 12-MAR-1999; 99US-0123951.
PR 28-MAY-1999; 99US-0136436.
PR 28-MAY-1999; 99US-0136437.
PR 28-MAY-1999; 99US-0136439.
PR 28-MAY-1999; 99US-0137127.
PR 28-MAY-1999; 99US-0137131.
PR 28-MAY-1999; 99US-0137567.
PR 30-JUN-1999; 99US-0141448.
PR 27-AUG-1999; 99US-0151114.
PR 03-SEP-1999; 99US-0152524.
PR 29-SEP-1999; 99US-0156633.
PR 29-SEP-1999; 99US-0156555.
PR 29-SEP-1999; 99US-0156634.
XX
XX (AREN-) ARENA PHARM INC.
XX
XX Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
XX
XX WPI; 2000-317986/27.
XX P-PSDB; AAB02838.

PT Non-endogenous, human G protein-coupled receptors for screening

XX receptor, inverse or partial agonists useful as therapeutic agents -

XX Example 1; Page 105-106; 187pp; English.

XX The present invention describes transmembrane receptors, preferably
CC human G protein coupled receptors (GPCR), for which the endogenous
CC ligand is unknown (orphan GPCR receptors). More specifically the present
CC invention relates to non-endogenous, constitutively activated versions
CC of a human GPCR. These non-endogenous human GPCRs can be useful for
CC the direct identification of candidate compounds as receptors agonists,
CC inverse agonists or partial agonists for use as pharmaceutical agents.
CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
CC the exemplification of the present invention.

XX Sequence 1080 BP; 162 A; 369 C; 322 G; 227 T; 0 other;

Alignment Scores:

Pred. No.: 1.25e-36 Length: 1080
Score: 494.00 Matches: 106
Percent Similarity: 53.54% Conservative: 53
Best Local Similarity: 35.69% Mismatches: 116
Query Match: 23.13% Indels: 22
DB: 21 Gaps: 6

US-09-208-629F-3 (1-407) x AAA46032 (1-1080)

Qy 96 AsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIle 115

Db 28 GACAAGCGGACGCTGCAGATGCTGCGGAACCCCGCGATCGCGTGGCCGCGGTG 87

Qy 116 TyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeuSer 135

Db 88 TACTCGCTGGTGGCGGTCAGCATCCCGGGCAACCTCTCTCTGTGGTGTGTGC 147

Qy 136 LeuArgThrLysSerIleSer---LeuValIlePheHisThrAsnLeuAlaIleAlaAsp 154

Db 148 CGCGCGATGGGCGCCAGATCCCGTGGTCACTTCATGATCAACCTGACCGCGACG 207

Qy 155 LeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyr 174

Db 208 CTGATCGTGGCCAGCGTGTGCTTTCCAAATCTACTACCATTCGCAACCCACCACCTGG 267

Qy 175 ValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCys 194

Db 268 GTATTGGGGTGGCTTTCACACCTGGTGGACCGTGGCCCTTTTACGCCAACATGTATTCC 327
 Qy 195 AlaIleuLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPhe 214
 Db 328 AGCATCTCACCAGCTGATCAGCTGGAGCGCTTCTGGGGTCTGTACCGCTC 387
 Qy 215 ThrTyrGlnIleuLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpVal 234
 Db 388 AGCTCCAAAGCGTGGCCCGCTGTTACGGCGTGGCCGCTGTGAGGAGCCTGGCTG 447
 Qy 235 MetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHis 254
 Db 448 CTGCTCTCAGCGCCCTGTGGCCCGTGGCGGCACCGCATCTCACCTACCGGTGCAGCC 507
 Qy 255 SerGluLeuThrCysHisAspValValAspAlaCysGluSerProSer----- 271
 Db 508 CTGGGCATCATCAGCTGCTCGAGCTCTCAAGTGGAGCATGTCTCCCGCAGGTGGCCATG 567
 Qy 272 -----SerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuLeuPro 289
 Db 568 TGGCGCGTGTCTCTTCCATCTTCCATCTTCCATCTGCTG-----TTCTCATCCCG 615
 Qy 290 PheValIleIlePheCysTyrThrThrLeuLeuHisLysLeu----- 304
 Db 616 TTCTGATCAGCGTGGCTGTTACCGCCACCATCTCAAGCTGTTCGCGCAGGAGGAG 675
 Qy 305 -----LysSerLysAspArgIleTyrIleuGlyTyrIleLysAlaValLeuLeuLeu 321
 Db 676 GCGCACGGCGGAGCAGCGAGCGCGGTGGCGCTGGCGCGGTGCTTCTG--- 732
 Qy 322 LeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAla 341
 Db 733 -----GCCTTTGTACCTGCTGCTCGCCCAACAACTTCTGCTCTCGCGGCACATCGTG 786
 Qy 342 AsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeu 361
 Db 787 AGCCGCTGTTCACGCG---AAGAGCTACTACACGCTGTACAGCTCACGCTGTGTCTC 843
 Qy 362 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
 Db 844 AGCTGCTCAACAACTGCTGGACCGCTTGTATTATTACTTTGGTCTCCCG 894
 RESULT 38
 ID AAX90979 standard; cDNA; 1361 BP.
 XX
 AC AAX90979;
 XX
 DT 25-JAN-2000 (first entry)
 XX
 DE Mouse protease-activated receptor 4 (PAR4) cDNA.
 XX
 KW Mouse protease-activated receptor 4 cDNA; PAR4; thrombin receptor;
 KW phosphoinositide hydrolysis; calcium efflux; platelet aggregation;
 KW platelet activation; thrombin agonist; thrombin antagonist; therapeutic;
 KW wound; blood coagulation; heart attack; stroke; inflammatory response;
 KW proliferative response; atherosclerosis; restenosis; glomerulosclerosis;
 KW pulmonary inflammation; ARDS; adult respiratory distress syndrome; ds.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT 1..1191
 FT CDS /*tag= a
 FT /product= "PAR4"
 XX
 FN W09943809-A2.
 XX
 PD 02-SEP-1999.
 XX
 PF 11-FEB-1999; 99WO-US02983.
 XX
 PR 27-FEB-1998; 98US-0032397.

XX (REGC) UNIV CALIFORNIA.
 PA Coughlin SR, Kahn M;
 XX WPI; 1999-619953/53.
 DR P-PSDB; AAY15081.
 XX
 PT DNA molecules encoding protease-activated receptor 4, useful in
 compound assays for thrombin agonist and antagonist activity -
 XX
 PS Claim 3; Fig 1; 69pp; English.
 XX
 CC The present sequence is a cDNA encoding protease-activated receptor 4
 CC (PAR4) and derived from a mouse 14-15 days embryo library. PAR4 is a G
 CC protein-coupled thrombin receptor expressed on cell surface. It is
 CC activated by thrombin and mediates signalling events e.g.
 CC phosphoinositide hydrolysis, calcium efflux and platelet aggregation.
 CC The receptor is highly expressed in spleen cells and likely to be
 CC involved in thrombin-mediated activation of platelets and other
 CC haematopoietic cells. It is used for screening novel thrombin agonists
 CC or antagonists. The agonists are used as therapeutics to treat wounds,
 CC promote clotting and as reagents to activate platelets in diagnostic
 CC tests. Antagonists are used to control blood coagulation, treat heart
 CC attacks and strokes, and block inflammatory and proliferative responses
 CC that occur in normal wound healing and variety of diseases including
 CC atherosclerosis, restenosis, pulmonary inflammation (ARDS) and
 CC glomerulosclerosis.
 XX
 SQ Sequence 1361 BP; 234 A; 416 C; 394 G; 317 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.69e-36 Length: 1361
 Score: 494.00 Matches: 105
 Percent Similarity: 53.03% Conservative: 70
 Best Local Similarity: 31.82% Mismatches: 139
 Query Match: 23.13% Indels: 16
 DB: 20 Gaps: 5
 US-09-208-629F-3 (1-407) x AAX90979 (1-1361)
 Qy 48 ProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPhePro 67
 Db 127 CCCACAGTAGAATCTCAAGAG-----CCGAGTCTCTCAGACAGCCTAATCCA 174
 Qy 68 LeuSerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGlu 87
 Db 175 -----CGAGGTACCGCGCAAAATTC-----TGTGCCAAC 204
 Qy 88 AspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 107
 Db 205 GACAGT---GACACGCTGGAGCTCCCGCCAGCTCTCAGACACTGTCTGGGGTGGGTA 261
 Qy 108 SerThrGlnValIleProIleThrIleLeuLeuPheValValGlyValProSerAsn 127
 Db 262 TCCACAGGCTGGTACCTGCTCTATGGCTGTGTGGCTGTGTGGCTGTGTGGCTGTGCAAT 321
 Qy 128 IleValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHis 147
 Db 322 GGGCTGGCGCTGTGGGTGTCTGGCCCAAGGGTGCACCGCTGCACCATCTTGCTC 381
 Qy 148 ThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyr 167
 Db 382 ACGAACCTGGCAGTGGCTGATTCGCTGTGGCCCTGTGGCCGCCACACGACTGGCTTAC 441
 Qy 168 HisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValVal 187
 Db 442 CACTTGGTGGCCAGCGCTGGCCATTTGGTAGAGTGTCTGGCGGTGGCCACAGCTGCC 501
 Qy 188 PheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyr 207
 Db 502 CTCATGGCCACATGTATGGTTAGTGTCTGGCTGGCAGTGCAGTGGACAGATAC 561

Qy 208 LeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeu 227
Db 562 CTGGCCCTGGTGCATCTTTGGGGCCGGTGGTGGTCAACGCCCTCACTACTGGA 621
Qy 228 MetCysGlyLeuValPheValMetValPheLeuTyrMetLeuProPheValLysLeuLys 247
Db 622 CTCGTGTTGGTGGCTGCTCTGCAGCCACCCCTGGCCCTTGCCTCTCACTCTGCATCGG 681
Qy 248 GlnCluTyrHisLeuValHisSerGluLeuThrThrCysHisAspValValAspAlaCys 267
Db 682 CAGACCTTCGATAGCTGGCTCCGATCGCATGCTGTGTATGATGCGTCCCTGACT 741
Qy 268 GluSerProSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeu 287
Db 742 GAGCAGACCTCCCACTGAGA---CCGGCTTTCATCTGCTGGCTGCTGGCTGCTTC 798
Qy 288 IleProPheValLysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLys 307
Db 799 GTGCCACTGCTGGCCATGGGCTGTGTATGGAGCCACCTTGTGTGCTGGGCGGCAAT 858
Qy 308 AspArgIleThrPheGlyTyrIleLysAlaValLeuLeuLeuLeuLeuValLysLeu 327
Db 859 GCGCAGCGCTACAGCCATGACTCAGACTGACAGCCCTGGTACTGTCTCGCAGTGGCC 918
Qy 328 CysPheAlaProThrAsnIleLeuValLysLeuLeuLeuLeuLeuLeuLeuLeuLeu 347
Db 919 TCCTTACACCTAGCATGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978
Qy 348 ThrAspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCys 367
Db 979 TGGGGCAATCTCTATGGAGCTATGTGCCAGCCCTGGCAGCTCAGCACCTCAACAGTGC 1038
Qy 368 LeuAspProPheLeuTyrPheValMetSer 377
Db 1039 GTAGACCTTTCATCTACTACTATGTGTC 1068

RESULT 39

AAAF57649
ID AAF57649 standard; cDNA; 1080 BP.
XX
AC AAF57649;
DT 29-JUN-2001 (first entry)
XX
DE Human G-protein coupled receptor, PAUL encoding cDNA.
XX
KW G-coupled protein receptor; PAUL; anti-HIV; antibacterial; antiviral;
KW antifungal; protozoacide; cytostatic; antidiabetic; anorectic; human;
KW antiasthmatic; antiparkinsonian; cardiant; hypertensive; osteopathic;
KW antianginal; cerebroprotective; antitumor; antidiabetic; antiepileptic;
KW tranquilizer; nootropic; anticonvulsant; neuroleptic; vaccine; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..1281
FT /tag= a
FT /product= "PAUL"
XX
FN WO200125280-A1.
XX
PD 12-APR-2001.
XX
PF 03-OCT-2000; 2000WO-US27228.
XX
PR 06-OCT-1999; 99US-0413534.
XX
PA (SMK) SMITHKLINE BEECHAM CORP.
XX
PI Zhu Y, Li X, Vawter L,
XX WPI; 2001-273570/28.
DR P-PSDB; AAB62285.
XX

XX Novel G-coupled protein receptor, PAUL useful for treating diseases
PT such as microbial infections, cancers, obesity, asthma, diabetes,
PT hypotension, osteoporosis, myocardial infarction, stroke, ulcer and
PT allergy -
XX
XX Claim 2; Page 25; 34pp; English.
XX
CC This cDNA encodes a human G-coupled protein receptor (designated PAUL).
CC The PAUL polypeptide and polynucleotide are useful for treating and
CC diagnosing infections such as bacterial, fungal, protozoan and viral
CC infections, particularly infections caused by human immunodeficiency
CC virus (HIV)-1 or HIV-2, pain, cancers, diabetes, obesity, anorexia,
CC bulimia, asthma, Parkinson's disease, acute heart failure, hypotension,
CC hypertension, urinary retention, osteoporosis, angina pectoris, stroke,
CC myocardial infarction, ulcers, allergies, benign prostatic hyperplasia,
CC migraine, vomiting, psychotic and neurological disorders including
CC anxiety, schizophrenia, manic depression, depression, delirium, dementia
CC and severe mental retardation, and dyskinesia such as Huntington's
CC disease or Gilles de la Tourette's syndrome. The PAUL sequences are
CC useful as vaccines to induce an immunological response in a mammal. The
CC PAUL polypeptide is also useful for identifying its modulators. The PAUL
CC polynucleotide is valuable for chromosome localization studies and for
CC tissue expression studies.
XX
SQ Sequence 1080 BP; 162 A; 370 C; 321 G; 227 T; 0 other;
XX

Alignment Scores:
Pred. No.: 1.54e-36 Length: 1080
Score: 493.00 Matches: 106
Percent Similarity: 53.54% Conservative: 53
Best Local Similarity: 35.69% Mismatches: 116
Query Match: 23.08% Indels: 22
DB: 22 Gaps: 6

US-09-208-629f-3 (1-407) x AAF57649 (1-1080)

Qy 96 AsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIle 115
Db 28 GACACCCGACGCTGCAGATGCTGGGAACCCGGGATCGGGTGGCCCTGCGGTGGTG 87
Qy 116 TyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSer 135
Db 88 TACTCGCTGGTGGCGGCTGAGATCCCGGGCAACCTCTCTCTCTGTGGGTGTGTC 147
Qy 136 LeuArgThrLysSerIleSer---LeuValIlePheHisThrAsnLeuAlaIleAlaAsp 154
Db 148 CGCGCGATGGGCGCCAGATCCCGCTGCTGCTCATCATCATCAACCTGACGCTACGGAC 207
Qy 155 LeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsp 174
Db 208 CTGATCGTGGCCAGCGGTGCTTTCCTTCCAAATCTACTACCATGCAACCGCCACTGG 267
Qy 175 ValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyrCys 194
Db 268 GTATTCGGGTGCTGCTTTGCAACGTTGGTGGCTGGCTTTTACGCAACATGATATCC 327
Qy 195 AlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPhe 214
Db 328 AGCATCTCACCATGACCTGTATCAGCGTGGAGCGCTTCTTGGGGTCTCTGTACCGCTC 387
Qy 215 ThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyrVal 234
Db 388 AGCTCCAAGCGTGGCGCGCGCTGCTTACGGGTGGCGCGCTGTGCAGGACCTGGCTG 447
Qy 235 MetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHis 254
Db 448 CTGCTCTGACCGCTGCTCCCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 507
Qy 255 SerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSer----- 271
Db 508 CTGGGATCATCATCTGCTTCCGACGCTCTCAAGTGGAGCATGTCTCCCGGCGGCGGCGG 567

```
QY 272 -----SerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuLeuPro 289
Db 568 TGGCGCGTGTCTCTTCCACCATCTCATCTGCTG-----TTCTCATCCCG 615
QY 290 PheValleileilePheCysTyrThrThrLeuLeuHisLysLeu----- 304
Db 616 TTCTGTATCACCCTGGCTTTTACACCGCCACCATCTCTCAAGCTGTGGCCACGGAGGAG 675
QY 305 -----LysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 321
Db 676 GCGCACGCGGAGGAGGAGGCGCGGTGGCGCTGGCGGTGGTCTTGTCTG--- 732
QY 322 LeuVallelePheThrIleCysPheAlaProThrAsnIleleleuValleleHisAla 341
Db 733 -----GCCTTTGTACCTCTTTCGCCCCCAACAACCTCTGTCTCTGGGGCACATGCTG 786
QY 342 AsnTyrTyrThrHisAsnThrAspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeu 361
Db 787 AGCGCGCTGTCTACGGC---AGAGTACTACTACACGTGTACAGCTTCACGCTGTGTCTC 843
QY 362 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 844 AGTGCCTCAACAACCTCTCTGGACCGCTTGTCTTATTACTTTGGCTCCCG 894

RESULT 40
ID AAI64231 standard; cDNA; 1080 BP.
XX
AC AAI64231;
XX
DT 12-MAR-2002 (first entry)
XX
DE G-protein coupled receptor (designated Paul) encoding sequence.
XX
KW G-protein coupled receptor; paul; 7TM receptor; GPC receptor;
KW human; antibacterial; viricide; fungicide; protozoicide;
KW anti-Human Immunodeficiency Virus; analgesic; cytosolic; antidiabetic;
KW anorectic; osteopathic; antiparkinsonian; cardiant; hypertensive;
KW hypotensive; osteopathic; antianalgesic; antianalgesic; anticonvulsant;
KW antitumor; antiallergic; antidepressant; antimigraine; antiemetic;
KW tranquilizer; neuroleptic; neuroprotective; neurotropic; anticonvulsant;
KW Human Immunodeficiency Virus type 1; HIV-1; HIV-2; pain; cancer;
KW diabetes; obesity; anorexia; bulimia; asthma; Parkinson's disease;
KW acute heart failure; hypertension; hypertensive; urinary retention;
KW osteoporosis; angina pectoris; myocardial infarction; stroke; ulcer;
KW allergy; benign prostatic hypertrophy; migraine; vomiting; anxiety;
KW schizophrenia; manic depression; delirium; dementia; mental retardation;
KW dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1080
FT /tag= a
FT /product= "Paul protein"
FT /note= "G-protein coupled receptor referred to as Paul
FT in the specification"
XX
XX
PN US2001029032-A1.
XX
PD 11-OCT-2001.
XX
PF 18-DEC-2000; 2000US-0739151.
XX
PP 09-OCT-1998; 98US-103789P.
PR 06-OCT-1999; 99US-0413534.
XX
PA (ZHUY/) ZHU Y.
PA (LIXX/) LI X.
PA (VAVT/) VAWTER L.
XX
PI Zhu Y, Li X, Vawter L;
XX
```

WPI; 2001-647985/74.
P-PSDB; AAG78530.

New G-protein coupled receptor polypeptide, referred as Paul and encoding polynucleotide, useful for diagnosing and treating cancers, infections, neurological disorders, diabetes, asthma and identifying modulators

Claim 2b; Page 12-13; 15pp; English.

The invention relates to an isolated polypeptide, a member of G-protein coupled receptor family of polypeptides, comprising a fully defined sequence of 359 amino acids, its 95% identical sequence, or a polypeptide encoded by a polynucleotide comprising a fully defined sequence of 1080 base pairs defined in the specification, or a fragment or variant of it. The activity of the protein of the invention may be described as CC antibacterial, viricide, fungicide, protozoicide, anti-Human Immunodeficiency Virus, analgesic, cytosolic, antidiabetic, anorectic, osteopathic, antiparkinsonian, cardiant, hypertensive, hypotensive, antiallergic, antidepressant, antimigraine, anticonvulsant, tranquilizer, neuroleptic, neuroprotective, neurotropic and anticonvulsant. Polypeptides and polynucleotides of the invention are useful in diagnosis and in identifying compounds such as agonists and antagonists which are useful in therapy. They are also useful for treating diseases, including infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by Human Immunodeficiency virus type 1 or 2 (HIV-1 or HIV-2), pain, cancers, diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers, allergies, benign prostatic hypertrophy, migraine, vomiting, psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, depression, delirium, dementia, and severe mental retardation, dyskinesias such as Huntington's disease or Gilles de la Tourette's syndrome. Polypeptides and polynucleotides of the invention may also be useful as diagnostic reagents, for example in the detection of mutations in the associated gene, chromosome localisation studies and expression pattern determination. The current sequence represents a G-protein coupled CC receptor (designated "Paul" in the specification) encoding sequence.

Sequence 1080 BP; 162 A; 370 C; 321 G; 227 T; 0 other;

Alignment Scores:

Pred. No.:	1,546-36	Length:	1080
Score:	493.00	Matches:	106
Percent Similarity:	53.54%	Conservative:	53
Best Local Similarity:	35.69%	Mismatches:	116
Query Match:	23.08%	Indels:	22
DB:	23	Gaps:	6

US-09-208-629F-3 (1-407) x AAI64231 (1-1080)

QY	96	AsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrClnValleProAlaIle	115
Db	28	GACACCGCAGCTGCAGATGCTCGCAACCCCGGATCGCGTGGCCCTCCCGCTGGTG	87
QY	116	TyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSer	135
Db	88	TACTCGCTGGTGGCGGCGTGCAGATCCCGGGCAACCTCTTCTCTGTGGGTGCTGTGC	147
QY	136	LeuArgThrLysSerIleSer---LeuValIlePheHisThrAsnLeuAlaIleAlaAsp	154
Db	148	CGGCGCATGGGCGCCAGATCCCGCTCGGTCACTTCATGATCACTGACGTCACGGAC	207
QY	155	LeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp	174
Db	208	CTGATGCTGCCAGCGTGTGCTTTTCAAAATCTACTACCATTCACACGCCACCACTGG	267
QY	175	ValPheGlyGluValMetCysArgIleThrValValPheTyrGlyAsnMetTyrCys	194
Db	268	GTATTCGGGTGCTGCTTTTGCAACAGTGGTGCACCGTGGCTTTTACGCAACATGATTTC	327

```
Qy 195 AlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProphe 214
Db 328 AGCATCTCTACCATGACTGTATCAGCGTGGAGCGCTTCTGGGGTCTCTGTACCCGCTC 387
Qy 215 ThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyrVal 234
Db 388 AGCTCCAAGCGCTGGCGCGCGCTGTATCGCGTGGCGCGTGTGCGAGGAGCTGGCTG 447
Qy 235 MetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHis 254
Db 448 CTGCTCTCGACCGCTGTCTCCCGCTGGCGCGCACCATCTCACCTACCCGGTGCAGGCC 507
Qy 255 SerGluIleThrThrCysHisAspValAlaAspAlaCysGluSerProSer----- 271
Db 508 CTGGGCATCATCCTCTTCGACGCTCTCAAGTGGAGCATGTCTCCAGCGTGGCCATG 567
Qy 272 -----SerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIlePro 289
Db 568 TGGGCGGTGTCTCTTCCATCTTCATCTCTGCTG-----TTCTCATCCCG 615
Qy 290 PheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeu----- 304
Db 616 TTCGTGATCACCGGTGTGTACACGCGCCACCATCTCAAGCTGTTCGCGACGGAGAG 675
Qy 305 -----LysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuIle 321
Db 676 CGGCACGCGCGGAGCGAGCGCGCGCTGGCGCTGGCGCGGTGGTCTTGTCTG--- 732
Qy 322 LeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAla 341
Db 733 -----GCCTTTGTACCTGCTTCGCCCCCAACAACTTCGTGCTCTTGGCGCACATCGTG 786
Qy 342 AsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeu 361
Db 787 AGCCGCTGTCTTCTACGGC---AAGAGCTACTTACCACGCTGTACAGCTCACGCTGTGTCTC 843
Qy 362 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 844 AGTGCCTCAACAACACTGTCTGGACCCGTTGTATTATTACTTTGTGCTCCCG 894
```

Search completed: June 29, 2003, 08:18:33
Job time : 243.264 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:29:54 ; Search time 2264.22 Seconds
(without alignments)
5231.316 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLYTXQHPVAGSQDKMKIL.....AMARPLXRRPRDIWIHAW 407

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09208629/runat_24062003_101625_12922/app_query.fasta_1.1165
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09208629 @CGN 1.1 2827 @runat_24062003_101625_12922 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pt.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_bu.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_mu.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

RESULT 1

ALIGNMENTS

Result No.	Score	Query Match %	Length	DB ID	Description
1	2121	99.3	1224	6	AR070424 Sequence
2	2121	99.3	2409	10	MMU92972 Mus musculus
C 3	1929	90.3	235097	2	AC110877 Mus musculus
4	1729.5	81.0	1505	10	AF310076 Rattus norvegicus
5	1443.5	67.6	1224	6	AR070426 Sequence
6	1443.5	67.6	1830	9	HSU92971 Homo sapiens
7	1372.5	64.3	9273	9	AF374726 Homo sapiens
C 8	1372.5	64.3	101887	9	AC026725 Homo sapiens
9	1372.5	64.3	131370	9	AC026706 Homo sapiens
10	1369	64.1	1102	6	AR070427 Sequence
11	1300.5	60.9	1124	6	AR070425 Sequence
12	584	27.3	2732	10	MPAR2MR Mus musculus
13	584	27.3	2732	6	AR012639 Sequence
14	584	27.3	2732	6	AR0171259 Sequence
15	584	27.3	2732	6	187850 Sequence 60
16	583	27.3	2772	10	BC025432 Mus musculus
17	582.5	27.3	1680	5	XLU09632 Xenopus laevis
18	581.5	27.2	1477	10	MMPRACR2 Mus musculus
C 19	581.5	27.2	199227	2	AC110876 Mus musculus
20	579.5	27.1	1475	6	AR012637 Sequence
21	579.5	27.1	1475	6	AR0171257 Sequence
22	579.5	27.1	1475	6	142454 Sequence 1
23	579.5	27.1	1475	6	187848 Sequence 1
24	569	26.6	2813	9	BC012453 Homo sapiens
25	568	26.6	2876	9	BC018130 Homo sapiens
C 26	564.5	26.4	1428	10	RNU61373 Rattus norvegicus
27	564.5	26.4	151166	2	AC130634 Rattus norvegicus
C 28	564	26.4	52368	9	AC010621 Homo sapiens
29	563	26.4	1255	6	AR012638 Sequence
30	563	26.4	1255	6	AR0171258 Sequence
31	563	26.4	1255	6	142455 Sequence 3
32	563	26.4	1255	6	187849 Sequence 3
33	563	26.4	1289	9	HSPAR2B Homo sapiens
34	563	26.4	1451	9	HSU34038 Homo sapiens
C 35	563	26.4	18351	9	AF400075 Homo sapiens
36	563	26.4	184536	2	AC068682 Homo sapiens
37	563	26.4	204485	2	AC114962 Homo sapiens
C 38	560.5	26.2	1124	9	HSU36753 Homo sapiens
39	559.5	26.2	88955	2	AC128225 Rattus norvegicus
C 40	551	25.8	1414	6	AR012640 Sequence
41	551	25.8	1414	6	AR0171260 Sequence
42	551	25.8	1414	6	187851 Sequence 62
43	540.5	25.3	3313	10	BC031516 Mus musculus
44	534.5	25.0	199227	2	AC110876 Mus musculus
45	533	25.0	5620	10	MMTHREC02 Mus musculus

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

AR070424
LOCUS AR070424 1224 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 1 from patent US 5892014.
ACCESSION AR070424
VERSION AR070424.1 GI:7221312
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1224)
AUTHORS Coughlin,S.R., Ishihara,H. and Connolly,A.
TITLE DNA encoding a protease-activated receptor 3
JOURNAL Patent: US 5892014-A 1.06-APR-1999;
FEATURES Location/Qualifiers
source
1. 1224
/organism="unknown"
BASE COUNT 299 a 341 c 239 g 345 t
ORIGIN
Alignment Scores:
Pred. No.: 2.6e-163 Length: 1224
Score: 2121.00 Matches: 404
Percent Similarity: 99.26% Conservative: 0
Best Local Similarity: 99.26% Mismatches: 3
Query Match: 99.30% Indels: 0
DB: 6 Gaps: 0
US-09-208-629F-3 (1-407) x AR070424 (1-1224)
Qy 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
Db 3 ACTTTGTATACTTAACAACATCCCTGAGCCGGGTCTCAGGACATCAAGATGAAATCCTT 62
Qy 21 IleLeuValAlaAlaGlyLeuPheLeuProValThrValCysGlnSerGlyIleAsn 40
Db 63 ATCTTGGTGTGAGCTGGCTGCTGTTCTGCCAGTCACCTGTTGCCAAAGTGGCATAAAT 122
Qy 41 ValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGln 60
Db 123 GTTTCAGACCACTCAGCAAGCCAACTTAACCTTAAGAGTTTATGGGGTCCCA 182
Qy 61 AsnThrPheGluGluPheProLeuSerAspIleGluGlyThrPheAlaThrThr 80
Db 183 AATACCTTTGAAGAATTCCTTACATAGAGGGCTGGACAGAGGCCACCACT 242
Qy 81 IleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIle 100
Db 243 ATAAAGCGGAGTGTCCGAGGAGCAGTATTTCAACTCTCCACGGAATAATGTACCATA 302
Qy 101 GlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPhe 120
Db 303 GGATACCTGAGAAGTCTCTTAAGTACCAAGTATACCTGCCATCTATATCTGCTGTT 362
Qy 121 ValValGlyValProSerAsnIleValThrLeuTyrLysLeuSerLeuArgThrLysSer 140
Db 363 GTGGTGTGTGTACCATCAACATCGTACCCCTGTGAAACTCTCTCTTAAGGACCAATCC 422
Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
Db 423 ATCAGTCTGGTCATCTTTTACCAACCACTGGCCATCGCAGATCTCTCTTCTGTGTGACA 482
Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
Db 483 CTGCCATTTAAGATCGCTACCATCTCAATGGCAACAACCTGGGTATTTGGGAGGTATG 542
Qy 181 CysArgIleThrThrValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
Db 543 TGGCGGATCAGCAGCGTGGTCTTCTACGGCAACATGTACTCGCTATCTCTCATCCTCACT 602
Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
Db 603 TGCATGGGCATCAACCGCTACTGGCCACGGCTCACCTTTCATACCAAGAGCTGCC 662

Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
Db 663 AAACGCAGCTTCTCTTGTCTCATGTGCGATAGTGTGGTTCATGGTTTCTTATACATG 722
Qy 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCys 260
Db 723 CTGCCCTTTGTCTCATCTGAAGCAGGAGTACCACTCGTCCACTCAGAGATCACCACTGC 782
Qy 261 HisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSer 280
Db 783 CACGATGTGTGACCGGTGGAGTCCCATCATCTCTCGGATCTTACTACTTCTGCTCC 842
Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
Db 843 TTAGCATTTCTTGGGTTCTCTCATCCGTTTGTGATCATCATCTTCTGTGTACAGACTCTC 902
Qy 301 IleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 320
Db 903 ATCCACAAACTTAAATCAAAGGATCGGATATGGCTGGCTACATCAAGGCGCTCTCCTC 962
Qy 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
Db 963 ATCTTGTGATTTTCAAAATTTGCTTTGCCCCCAACATCATACTCGTAATCCACCAT 1022
Qy 341 AlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCys 360
Db 1023 GCCAACTACTACTTACCACAAATACCGACAGCTTGATCTTTATGTATCTTATGCTCTGTC 1082
Qy 361 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValVal 380
Db 1083 CTGGGAGGCTGAATAGCTGCCTAGATCCATTCCTTTACTTGTCTCATGTCGAAGTTGTA 1142
Qy 381 AspGlnLeuAsnPro***SerAlaMetAlaArgProLeu***ArgProArgArgAspIle 400
Db 1143 GATCAGCTTAATCTTACTGCGCAATGGCAAGCACTTTAGAGACCAAGAGAGATATC 1202
Qy 401 TrpGluAspIleHisAlaTrp 407
Db 1203 TGGGAAGACATACATGCTTGG 1223
RESULT 2
MMU92972 2409 bp mRNA linear ROD 16-APR-1997
LOCUS Mus musculus protease-activated receptor 3 (PAR3) mRNA, complete
DEFINITION cds.
ACCESSION U92972
VERSION U92972.1 GI:1938376
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2409)
AUTHORS Ishihara,H., Connolly,A.J., Zeng,D., Kahn,M.L., Zheng,Y.W.,
Timmons,C., Tram,T. and Coughlin,S.R.
TITLE Protease-activated receptor 3 is a second thrombin receptor in
humans
JOURNAL Nature 386 (6624), 502-506 (1997)
MEDLINE 97242411
PUBMED 9087410
REFERENCE 2 (bases 1 to 2409)
AUTHORS Ishihara,H., Connolly,A.J., Zeng,D., Kahn,M.L., Zheng,Y.W.,
Timmons,C., Tram,T. and Coughlin,S.R.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-1997) CVRI, UCSF, 3rd and Parnassus, San
Francisco, CA 94143, USA
FEATURES Location/Qualifiers
source
1. .2409
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
1. .2409
/gene="PAR3"

CDS 153..1262
/gene="PAR3"
/note="thrombin receptor; coagulation protease"
/codon_start=1
/product="protease-activated receptor 3"
/protein_id="AAC53137.1"
/db_xref="GI:1938377"
/translation="MKLLIIVAAAGLLPLVTVCSGINVSDNSAKPTLTITKSNNGFGP
LFTEEPFLSDIEGWTGATTTIKAECPEDSISTIRVNNATIGYLRSLSTQVIPIAYIL
LPVGVPSNVTIWKLSLRKTSISLVIHTNLAIDLFCVTLPFKIAYHLNNGNWF
GEVNCRTITVVFVGNMYKAILITCGINRYLAHPFTYOKLPKPSFSLGCMGIVWV
MVFLYMLPPVLLQAEVHLVSEITCTCHDVVDACESPSFEYFVSLAEFGFLIPVI
IICVYTLHLKLSKRIWLGYIKAVULLIIVITFCAPNTIILVIHHANYHYHTDS
LYFWLIALCGLSLSNCLDPFLYFVMSKVVDQLNP"
BASE COUNT 598 a 652 c 454 g 703 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 5,55e-163 Length: 2409
Score: 2121.00 Matches: 404
Percent Similarity: 99.26% Conservative: 0
Best Local Similarity: 99.26% Mismatches: 3
Query Match: 99.30% Indels: 0
DB: 10 Gaps: 0

US-09-208-629F-3 (1-407) x MMU92972 (1-2409)

QY 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
DB 105 ACTTGTATTAACAAACATCTCTGTAGCGGGTCTCAGACATCAAGATGAAATCCTT 164

QY 21 IleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsn 40
DB 165 ATCTGGTTCAGCTGGGCTGCTTCTGCCAGTCACTGTTGCCAAAGTGGCATAAAT 224

QY 41 ValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyGlyProGln 60
DB 225 GTTTCAGCAACCTCAGCAAGCCCAACCTTAACATTAAGAGTTTTAATGGGGGTGCCAA 284

QY 61 AsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThr 80
DB 285 AATACCTTTGAAGATTTCCACATTTCTGACATAGAGGGCTGGACAGAGGCCACCAACT 344

QY 81 IleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIle 100
DB 345 ATAAAGCGGAGTCTCCGAGGACAGTATTTCAACTCTCCACGTGAATAATGCTACCAT 404

QY 101 GlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPhe 120
DB 405 GGATACCTGAGAAGTTCCTTAAGTACCAAGTGTATACCTGCTATATATCTGCTGTTT 464

QY 121 ValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLysSer 140
DB 465 GTGGTGGTGTACCATCAACATCGTGACCTGTGGAAACTCTCTTAAGAGCAACATCC 524

QY 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
DB 525 ATCAGTCTGTGTCATCTTTTACACCAACCTGGCCATCGCAGATCTCTCTTTCTGTGACA 584

QY 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
DB 585 CTGCATTTAAGATCGGCTACCATCTCAATGGCAACACTGGGTATTTGGCGAGGTCTATG 644

QY 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThr 200
DB 645 TGCCGGATACACCGTCGTTTCTACGGCAACATGACTGCGCTATCTGATCCTCACT 704

QY 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
DB 705 TGCATGGGCATCAACCGTACCTACCTGGCCACGGCTCACCTTTCATACACAGAGTGGCC 764

QY 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240

DB 765 AAACGAGCTTCTCCTGTCTCATGTGTGGCATAGTGTGGTCAATGGTTTCTTATACATG 824

QY 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
DB 825 CTGCCCTTTGTCTATCTGAAGCAGGAGTACCACCTCTGCTCCACTCAGAGATCACCACCTGC 884

QY 261 HisAspValValAspAlaCysGluSerProSerSerPheArgPheTyrTyrPheValSer 280
DB 885 CAGCATGTCTGACACGCTGCGAGTCCCATCATCTCTCCGATTTCTACTACTTCTGCTCC 944

QY 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
DB 945 TTAGCATCTCTTGGGTCTCTCATCCCTTTGTGATCATCATCTTCTGTACAGACTCTC 1004

QY 301 IleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeu 320
DB 1005 ATCCACAACTTAATCAAGGATCGATATGCTGGCTACATCAAGGCCGTCTCTCTC 1064

QY 321 IleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisHis 340
DB 1065 ATCTTGTGATTTTCAAAATTTGCTTTGCCCCCAACATCATCTCGTAAATCCACCAT 1124

QY 341 AlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleIleAlaLeuCys 360
DB 1125 GCCAACTACTACTACCAATACCGACGCTTGTACTTATGTATCTTATGTCTGTGTC 1184

QY 361 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValVal 380
DB 1185 CTGGGAGCTGATAGTGCCTAGATCCATCTCTTACTTGTCTGTCGAAAGTTGTA 1244

QY 381 AspGlnLeuAsnPro***SerAlaMetAlaArgProLeu***ArgProArgArgAspIle 400
DB 1245 GATCAGCTTAATCTTCTAGTCGCAATGGCAAGACCCTTTAGAGACCAAGGAGATATC 1304

QY 401 TrpGluAspIleHisAlaTrp 407
DB 1305 TGGGAAGACATACATGCTTGG 1325

RESULT 3
AC110877 235097 bp DNA linear HTG 16-FEB-2002
LOCUS Mus musculus clone RP23-345121, WORKING DRAFT SEQUENCE, 3 unordered
DEFINITION pieces.
AC110877
AC110877.1 GI:18693419
VERSION HTG: HTGS PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 235097)
AUTHORS Martin, J., Hosseini, R., Peng, Y., Peng, Z., Rubin, E.M. and Cheng, J.F.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 235097)
AUTHORS Martin, J., Hosseini, R., Peng, Y., Peng, Z., Rubin, E.M. and
Cheng, J.-F.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2002) Berkeley PCA, Lawrence Berkeley National
Laboratory, 1 Cyclotron rd., Berkeley, CA 94720, United States of
America
COMMENT Draft Sequence Produced by Berkeley PGA
pgs.lbl.gov
Center Code: PCABERK.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2480: contig of 2480 bp in length

* 2481 2580: gap of unknown length
 * 2581 100301: contig of 97721 bp in length
 * 100302 100401: gap of unknown length
 * 100402 235097: contig of 134696 bp in length.

FEATURES
 source
 1. 235097
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP23-345121"

BASE COUNT 64117 a 51797 c 51868 g 67115 t 200 others
 ORIGIN

Alignment Scores:
 Pred. No.: 3,61e-145 Length: 235097
 Score: 1929.00 Matches: 365
 Percent Similarity: 97.63% Conservative: 5
 Best Local Similarity: 96.31% Mismatches: 5
 Query Match: 90.31% Indels: 4
 DB: 2 Gaps: 1

US-09-208-629F-3 (1-407) x AC110877 (1-235097)

Qy 29 PheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSerAlaLysPro 48
 Db 228815 TTCTTTCAATTACA-----GGCATAAATGTTTCAGACAACCTCAGCAAGCCCA 228768
 Qy 49 ThrLeuThrIleLysSerPheAsnGlyProGlnAsnThrPheGluGluPheProLeu 68
 Db 228767 ACCTTAACATTAAGAGTTTAAATGGGGGTCCCAAAATACCTTTGAAGAATTTCCCACTT 228708
 Qy 69 SerAspIleGluGlyTyrThrGlyValThrThrIleLysAlaGluCysProGluAsp 88
 Db 228707 TCTGACATAGAGGCTGACAGGAGCCACCAACTATAAAGCGAGTGTCCGAGGAC 228648
 Qy 89 SerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSer 108
 Db 228647 AGTATTTCAACTCTCCAGGTGAATAATGCTACCATAGATACCTTGAGAGTTCCTTAAGT 228588
 Qy 109 ThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIle 128
 Db 228587 ACCCAAGTGATACCTGCGCATCTACATCTGCTGTTTGTGGTGTTGACGACCAACATC 228528
 Qy 129 ValThrLeuTyrLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThr 148
 Db 228527 GTGACCTGTGAAACCTCTCTTAAAGACCAATCCATCAGTCTGTCATCTTTACACC 228468
 Qy 149 AsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHis 168
 Db 228467 AACCTGGCATCGCAGATCTCTTTCTGTGTACACACTGGCCATTTAAGATCGCCTACCAT 228408
 Qy 169 LeuAsnGlyAsnAsnTyrValPheGlyGluValMetCysArgIleThrThrValValPhe 188
 Db 228407 CTCATGGCAACAACCTGGTATTTGGCGAGGTACAGTCCCGGATCACACCGTGTGTTTC 228348
 Qy 189 TyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeu 208
 Db 228347 TACGGCAACATGACTCGGTATCTCTGATCTCACTGATCGGATCACACCGTACTCTG 228288
 Qy 209 AlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMet 228
 Db 228287 GCCACGGCTCACCTTTTCCATACACAGAGTGGCCCAACAGCAGCTTCTCCATGCTCATG 228228
 Qy 229 CysGlyIleValTyrValMetValPheLeuTyrMetLeuProPheValIleLeuLysGln 248
 Db 228227 TGTGGCATGGTGGGTGATGGTTTCTTATACATGCTGGCCTTTGTCACTCCCTGAACGAC 228168
 Qy 249 GluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGlu 268
 Db 228167 GAGTACCACTCTGCTCCACTCCAGATCACCACTGCCAGCATGCTGTTGACGGTGGCAG 228108
 Qy 269 SerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIle 288
 Db 228107 TCCCATCATCTCTCCGATTTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 228048

Qy 289 ProPheValIleIleIlePheCysTyrThrThrLeuIleHisIleLysLeuLysSerLysAsp 308
 Db 228047 CCGTTTGTGATCATCATCTCTGTTACAGACTCTCATCCAACTTAAATCAAGGAT 227988
 Qy 309 ArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCys 328
 Db 227987 CGGATATGCTGGGTACATCAAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 227928
 Qy 329 PheAlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThr 348
 Db 227927 TTTGCCCCCAACCAATCATCTCGTAATCCACCTGCACTACTACTACCAACAATACC 227868
 Qy 349 AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
 Db 227867 GACAGCTTGTACTTTATGATATCTTATGCTCTGTGCTGGGAGCTGATAGTCTGCTA 227808
 Qy 369 AspProPheLeuTyrPheValMetSerLysValValAspGlnLeuAsnPro***SerAla 388
 Db 227807 GATCCATCTCTTACTTTGCTCATGCGAAAGTTGTAGATCAGCTTAATCTTAGTCAGCA 227748
 Qy 389 MetAlaArgProLeu***ArgProArgArgAspIleTyrGluAspIleHisAlaTyrP 407
 Db 227747 ATGGCAAGACCACTTAGAGACCAAGAGAGATATCTGGGAAGACGATCATGCTTGG 227691
 RESULT 4
 LOCUS AF310076 1505 bp mRNA linear ROD 02-NOV-2001
 DEFINITION Rattus norvegicus protease activated receptor 3 mRNA, complete cds.
 ACCESSION AF310076
 VERSION AF310076.1 GI:16588398
 KEYWORDS Rattus norvegicus.
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1505)
 AUTHORS Chien,E.K., Marietti,S., Mendoza,J. and Phillippe,M.
 TITLE Cloning of the rat protease activated receptor isoforms 3 and 4
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1505)
 AUTHORS Chien,E.K., Marietti,S., Mendoza,J. and Phillippe,M.
 TITLE Direct Submission
 JOURNAL Submitted (02-OCT-2000) Obstetrics and Gynecology, University of
 Chicago, 5841 South Maryland Ave, MC2050, Chicago, IL 60637, USA
 FEATURES
 Location/Qualifiers
 1..1505
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="spleen"
 63..1169
 /note="PAR3; G-protein coupled receptor"
 /codon_start=1
 /product="protease activated receptor 3"
 /protein_id="AAL26789.1"
 /db_xref="GI:16588399"
 CDS
 1..1505
 /translation="MEMKVLILVGVRLFLPTTVCSGMKHSVSDNSALTAESFNGNEH
 SFEFPLSDIEGTGATTTKAKPEESITTLHVNNTATGYSLSSTKVIPIAYILV
 FIVGPANITLWKLSSRTKSLCVIFHTNLALIDLLFCVTLFPFKIAYHLNGNDWVFG
 EVMCRVTTVAFGNMYCAILLTGMINGRYLATHPFTYKLPKRNFTLLMCGVWM
 VVLMPLAILKQEVHLPQGITTCVDHDTCEPLPFQFYFVSLAFGFLIPFVVS
 VCYTLTHKLNQDRKWRKYKAVLLILVITICFAPTNILIIHHANYYSNTDSL
 YFMYLALCGLSLSCLDPFLFIMSKVDQLTS"
 BASE COUNT 359 a 394 c 316 g 436 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,96e-131 Length: 1505
 Score: 1729.50 Matches: 328
 Percent Similarity: 89.24% Conservative: 37
 Best Local Similarity: 80.20% Mismatches: 38

Query Match: 80.97% Indels: 6
DB: 10 Gaps: 3
US-09-208-629F-3 (1-407) x AF310076 (1-1505)

Qy 1 ThrLeuTyrThr***GlnHisProValAlaGlySerGlnAspIleLysMetLysIleLeu 20
Db 21 TCTCTGCACACTTAGTGACATCCCATAAACCGGGTCTCAAGACATGAGATGAAGTCTT 80
Qy 21 IleLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIle--- 39
Db 81 ATCTTGGTGGGGTAGCTGCTGTTCTGCCGACCACTGTTTCCAAAGTGGCATGAAA 140
Qy 40 AsnValSerAspAsnSerAlaLysProThrLeuThrIleLysSerPheAsnGlyPro 59
Db 141 CATGTTTCAGACAACTCAGCC-----TTAACTGCTGAGAGCTTTTAATGGCAAC-- 188
Qy 60 GlnAsnThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThr 79
Db 189 GAACATTCTCTTGAAGAAATCCCACTTTCTGACATAGAGGCTGGACAGGACCCACACA 248
Qy 80 ThrIleLysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThr 99
Db 249 ACTATAAAGCGAAGTCTCCGAGAAAGCAATACAACTCTCCATGTGAATAATGCTACC 308
Qy 100 IleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeu 119
Db 309 ATGGGATACCTGAGAAGTCTCTTAAGTACCAAGATGATACCTGCCATCATCTCGTGGT 368
Qy 120 PheValValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThrLys 139
Db 369 TTTGTGATGGTGTAACGAGCAACATCGTCACTGTGGAACCTCTCTCAAGGACCAAA 428
Qy 140 SerIleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysVal 159
Db 429 TCCATCTGTCTGGTCATCTTTTACCAACCACTGGCCATGCGGATCTCTTTCTGTGTC 488
Qy 160 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 179
Db 489 ACGTGTGGTCTTAAGATCGCTACCACTCAATGGGAACGACTGGGTCTTTGGCGAGTC 548
Qy 180 MetCysArgIleThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeu 199
Db 549 ATGTGCGGGGTACACGAGTCTTCTACGGCAACATGATTTGTGTATTTCTCATCTC 608
Qy 200 ThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeu 219
Db 609 ACCTGTATGGCATCAACCGCTACTCTGGCACGGTCCATCTCCCTTCACATACCGCAAGCTG 668
Qy 220 ProLysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyr 239
Db 669 CCCAAACGCAACTTCATTTGCTCATGTGTGGCGTGGTGGTGCATGGTGTCTTATAT 728
Qy 240 MetLeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThr 259
Db 729 ATGTGTGGCTTGGCCATCTCTCAAGCAGGAGTACCATCTTGTCCAAACAGGAGTACCAC 788
Qy 260 CysHisAspValValAspAlaCysGluSerProSerPheArgPheThrTyrPheVal 279
Db 789 TGGCACAGCTTCCACGACACATGGAGTCCCGTGGCTTCCAGTCTTACTACTTCTGTC 848
Qy 280 SerLeuAlaPhePheGlyPheLeuIleProPheValIleIlePheCysTyrThr 299
Db 849 TCCTTGGCTCTTTGGGTCTTATCCCATTTGGTGCAGTGTCTTCTGTACAGACT 908
Qy 300 LeuIleHisLysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeu 319
Db 909 CTCATCACAAAGCTTAACGCGCAGGATCGCAAAATGGGTGAGGTACATCAAGGCGTCTC 968
Qy 320 LeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHis 339
Db 969 CTCATCTTGTGATTTTTCACCATCTGCTTTGGCCCCCAACCATCATCATTAATTCAC 1028

340 HisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeu 359
1029 CACGCCAACTACTACTACAGCAACACTGATAGCTTGTACTTATGTATCTCATAGCTCTC 1088
Qy 360 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLysVal 379
1089 TGCTGGGAGCGCTGAATAGTTCCTAGACCCCAATCTCTTATTTATCATGTCAAAAT 1148
Qy 380 ValAspGlnLeuAsnPro***SerAlaMetAlaArgProLeu***ArgProArgAsp 399
1149 GTAGATCAGCTGACCTCGTAGTCAACAATGGCGAGATCATCTTGGAGACAAGATAGAT 1208
Qy 400 IleTrp-GluAspIleHisAlaTrp 407
1209 ATCTGGGGAATGTATATGCTTGG 1233

RESULT 5
AR070426
LOCUS AR070426 1224 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 4 from patent US 5892014.
ACCESSION AR070426
VERSION AR070426.1 GI:7221314
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE 1 (bases 1 to 1224)
AUTHORS Coughlin, S.R., Ishihara, H. and Connolly, A.
TITLE DNA encoding a protease-activated receptor 3
JOURNAL Patent: US 5892014-A 4 06-APR-1999;
FEATURES
Location/Qualifiers
1..1224
source
BASE COUNT 303 a 316 c 221 g 384 t
ORIGIN
/organism="unknown"

Alignment Scores:
Pred. No.: 2,54e-108 Length: 1224
Score: 1443.50 Matches: 264
Percent Similarity: 83.74% Conservative: 45
Best Local Similarity: 71.54% Mismatches: 59
Query Match: 67.58% Indels: 1
DB: 6 Gaps: 1

US-09-208-629F-3 (1-407) x AR070426 (1-1224)

Qy 11 GlySerGlnAspIleLysMetLysIleLeuIleLeuValAlaAlaGlyLeuLeuPheLeu 30
Db 40 GGGACTCAGCTCATCAAAATGAAAGCCCTCATCTTTGCCAGCTGCTGGCTCTCTCTG 99
Qy 31 ProValThrValCysGlnSerGlyIle---AsnValSerAspAsnSerAlaLysProThr 49
Db 100 TTGCCCACTTTTGTCTCAGAGTGGCATGGAAATGATACAAACACTTGGCAAGCCCAACC 159
Qy 50 LeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSer 69
Db 160 TTACCATTAAGACCTTTCGTGGAGCTCCGCCAAATCTTTTGAAGAGTTCCTCTTCT 219
Qy 70 AspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSer 89
Db 220 GCCTTGAAGGCTGGACAGGAGCCAGGATTTACTGTAAAAAATTAAGTGCCTCGAAGAA 279
Qy 90 IleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThr 109
Db 280 GCTTTCATCTCATGTGAAAAATGTACCATGGGTACCTGACCAGCTCTTAAAGTACT 339
Qy 110 GlnValIleProAlaIleTyrIleLeuPheValIleValGlyValProSerAsnIleVal 129
Db 340 AAATGATACCTGCCATCTACCTCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 399
Qy 130 ThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThrAsn 149
Db 400 ACCCTGTGGATGCTTTTCTTTCAGGACGAGATCCATCTGTACCACTGTGATTTCTACCA 459

Db 6276 TGTGGACTGGTGGGCAACAGTTTTCTATATATGCTGCCATTTTTCATCTAAGCAG 6335
Qy 249 GluTyrHisLeuValHisSerGluLeuThrThrCysHisAspValValAspAlaCysGlu 268
Db 6336 GAATATTATCTTTTCAGCCGACATCACCACCTGCCATGATGTTTCAACACACTTGGAG 6395
Qy 269 SerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeu 288
Db 6396 TCCTCATCTCCCTTCCAACTCTATTACTTCTCTTGGCATTTCTTGGATTCTTAAT 6455
Qy 289 ProPheValLeuLeuLeuPheCysTyrThrThrLeuLeuHisLysSerLysAsp 308
Db 6456 CCATTGTGCTTATCACTACTGCTATGACGACATCCTGCGACACTAATGATAGAT 6515
Qy 309 ArgLeuTrpLeuGlyTyrLeuLeuAlaValLeuLeuLeuValLeuPheThrLeuCys 328
Db 6516 CATAGATGTTGTGTATGTTAAGCGAGTCTCTCATCTTGTGATTTTACCATTTGC 6575
Qy 329 PheAlaProThrAsnLeuLeuValLeuValHisAlaLeuTyrTyrHisAsnThr 348
Db 6576 TTTGCTCCAAGCAATATTATTCTTATTATTCACCATGCTAATACTACTACTACAACACT 6635
Qy 349 AspSerLeuTyrPheMetTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 6636 GATGGCTTATTTTATATATCTCATAGCTTTGTGCTGGTAGTCTTAATAGTTGCTTA 6695
Qy 369 AspProPheLeuTyrPheValMetSerLys 378
Db 6696 GATCATCTCTTATTTTTCATGTCATAA 6725

RESULT 8
AC026725/c
LOCUS 101887 bp DNA linear PRI 21-JUL-2001
DEFINITION Homo sapiens chromosome 5 clone CTD-2236F14, complete sequence.
ACCESSION AC026725
VERSION AC026725.7 GI:14993696
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 101887)
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE 2 (bases 1 to 101887)
TITLE DOE Joint Genome Institute.
AUTHORS DOE Joint Genome Institute.
JOURNAL Direct Submission
REFERENCE 3 (bases 1 to 101887)
TITLE Submitted (23-MAR-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE 4 (bases 1 to 101887)
TITLE Submitted (27-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
REFERENCE 5 (bases 1 to 101887)
TITLE Submitted (21-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
COMMENT On Jul 21, 2001 this sequence version replaced gi:14572130.
Draft sequence produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.1.
STS Content:
SHGC-102125 G56565
SHGC-101620 G56147
SHGC-83998 G53567.
Location/Qualifiers
1. .101887

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2236F14"
BASE COUNT 29200 a 20058 c 21105 g 31524 t
ORIGIN
Alignment Scores: 2.09e-100 Length: 101887
Pred. No.: 1372.50 Matches: 248
Score: 83.71% Conservative: 45
Percent Similarity: 70.86% Mismatches: 54
Best Local Similarity: 64.26% Indels: 3
Query Match: 9 Gaps: 1
DB:
US-09-208-629F-3 (1-407) x AC026725 (1-101887)
Qy 29 PheLeuProValThrValCysGlnSerGlyLeuAsnValSerAspAsnSerAlaLysPro 48
Db 93492 TTTCTTTCAATTACAGGCATGGAA-----AATGATACAAACAACCTTGGCAAGCCA 93442
Qy 49 ThrLeuThrLeuLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 93441 ACCTTACCATTAAGACCTTTCTGTGGAGCTCCCCAAATTCCTTTGAAGAGTTTCCCTTT 93382
Qy 69 SerAspIleGluGlyTyrThrGlyAlaThrThrIleLysAlaGluCysProGluAsp 88
Db 93381 TCTGCTTGGAGGCTGGACAGGACGACGATTACTGTAAATTAAGTCCCTGAAGAA 93322
Qy 89 SerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSer 108
Db 93321 AGTCTTTCATCTCCATGTGAAATATGCTTACCATGGGTACCTGACCACTCTTAAAT 93262
Qy 109 ThrGlnValIleProAlaIleTyrIleLeuLeuPheValGlyValProSerAsnIle 128
Db 93261 ACTAAACTGATACCTGCCATCTACCTCTGGTGTGTAGTTGGTCTCCCGGCAATGCT 93202
Qy 129 ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThr 148
Db 93201 GTGACCTCTGGATGCTTTCTTCAGACCAAGATCCATCTGTACCACTGTTATCTACAC 93142
Qy 149 AsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHis 168
Db 93141 AACCTGCCATTGCAGATTTCTTTTGTGTATACATTCCTTTTAAAGTAGCTTATCAT 93082
Qy 169 LeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPhe 188
Db 93081 CTCAATGGGAACAACCTGGGTATTTGGAGAGGTCTGTGCGGCGCCACCACTCATCTTC 93022
Qy 189 TyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeu 208
Db 93021 TATGGCAACATGATCTGCTCCATTTCTGCTTGTGCTGTACATCAGATCAACCTTACCTG 92962
Qy 209 AlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMet 228
Db 92961 GCCATCGTCCATCTTTTACCTACCGGGCTGCCCAAGCACACCTATATGCTTGGTAACA 92902
Qy 229 CysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGln 248
Db 92901 TGTGGACTGGTGGGCAACAGTTTCTTATATATGCTGCCATTTTTCATCTAAGAACAG 92842
Qy 249 GluTyrHisLeuValHisSerGluLeuThrThrCysHisAspValValAspAlaCysGlu 268
Db 92841 GAATATATCTTGTTCAGCCAGACATCACCCTGCCATGATGTTTCAACACACTTGGCAG 92782
Qy 269 SerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeu 288
Db 92781 TCCTCATCTCCCTTCCAACTCTATTACTTCTCTCTGTCATCTCTTGGATTCTTAAAT 92722
Qy 289 ProPheValIleIleIlePheCysTyrThrThrLeuLeuHisLysLysSerLysAsp 308
Db 92721 CCATTTGTGCTTATCTACTGCTATGACGACCATCATCCGGACACTTAATGATACGAT 92662


```

Qy 309 ArgileTrpLeuGlyTyrIleLeuValLeuLeuLeuValIlePheThrIleCys 328
Db 92661 CATAGATGGTGTGGTATGTAAGCGAGCTCTCTCATCTTGTGATTTTACCAATTGC 92602
Qy 329 PheAlaProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrHisAsnThr 348
Db 92601 TTTGCTCCAGCAATATTATTCTTATTATTACCATTGCTAACTACTACAACCACT 92542
Qy 349 AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 92541 GATGGCTTATATTTTATATATCTCATAGCTTTGTGCTGGGTAGTCTTAATAGTTGCTTA 92482
Qy 369 AspProPheLeuTyrPheValMetSerLys 378
Db 92481 GATCCATTCTTTTATTTCTCATGTCAAAA 92452

RESULT 9
AC026706 131370 bp DNA linear PRI 21-JUN-2001
LOCUS Homo sapiens chromosome 5 clone CTD-2068N7, complete sequence.
DEFINITION AC026706
ACCESSION AC026706
VERSION AC026706.5 GI:14518404
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 131370)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished
2 (bases 1 to 131370)
DOE Joint Genome Institute.
Direct Submission
Submitted (23-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 131370)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (21-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jun 21, 2001 this sequence version replaced gi:13677033.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.3.
STS Content:
WI-11362 G24506.
FEATURES
Location/Qualifiers
1..131370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2068N7"
BASE COUNT 40462 a 26452 c 25592 g 38854 t
ORIGIN
Alignment Scores:
Pred. No.: 2,78e-100 Length: 131370
Score: 1372.50 Matches: 248
Percent Similarity: 83.71% Conservative: 45
Best Local Similarity: 70.86% Mismatches: 54
Query Match: 64.26% Indels: 3
DB: 9 Gaps: 1
US-09-208-629f-3 (1-407) x AC026706 (1-131370)
Qy 29 PheLeuProValThrValCysGlnSerGlyIleAsnValSerAspAsnSerAlaLysPro 48
Db 60186 TTCTTTCAATTACAGGATGAA-----AATGATACAAACAACTTGGCAAAGCCA 60236

```

```

Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 60237 ACCTTACCCATTAGACCTTTTCGTGGAGCTCCCCCAATCTTTTGAAGAGTTCCCTTT 60296
Qy 69 SerAspIleGluGlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAsp 88
Db 60297 TCTGCCTTGAAGGCTGGACAGGAGCCACGATTACTGTAAAAAATTAAGTGCCTGAAGAA 60356
Qy 89 SerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeu 108
Db 60357 AGTGCCTTCACATCTCCATGTGAAAAATGTACCAGGGTACCTGACCAGCTCCTTAAGT 60416
Qy 109 ThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIle 128
Db 60417 ACTAACTGATACCTGCCATCTACCTCCTGGTGTGTAGTTGGTGTCCCGGCCAATGCT 60476
Qy 129 ValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePheHisThr 148
Db 60477 GTGACCTGTGGATGCTTTTCTTCAGGACGACATCCATCTGTACCACCTGTATTCTACACC 60536
Qy 149 AsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHis 168
Db 60537 AACCTGGCCATTGCAGATTCTTTTGTGTGTACATTGCCCTTTAAGATAGATTATCAT 60596
Qy 169 LeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValPhe 188
Db 60597 CTCAAATGGGNACAACTGGGTATTTGGAGAGGCTCTGTGCGGGCCACCACAGTCACTTC 60656
Qy 189 TyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyrLeu 208
Db 60657 TATGGCAACATGACTGCTCCATTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60716
Qy 209 AlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMet 228
Db 60717 GCATCGTCCATCTTTCACCTACCGGGGCTGCCCAAGCACACCTATGCTTGGTGAACA 60776
Qy 229 CysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGln 248
Db 60777 TGTGGACTGGTGGGCAACAGTTTCTTATATATGCTGCCATTTTTCATCTGAGGACAG 60836
Qy 249 GluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGlu 268
Db 60837 GAATATTATCTTGTTCAGCCAGACATCACCCCTGCCATGATGTTTCAACACACTTGGCAG 60896
Qy 269 SerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPheGlyPheLeuIle 288
Db 60897 TCCTCATCTCCCTCCCAACTTATTACTCTCTCTGTCATCTCTGTCATCTTGGATTTCTAAT 60956
Qy 289 ProPheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerLysAsp 308
Db 60957 CCATTTGTGCTTATCATCTACTGCTATGCGAGCCATCATCGGACACTTAATGATACAGAT 61016
Qy 309 ArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeuValIlePheThrIleCys 328
Db 61017 CATAGATGGTGTGGTATGTTAAGGCGAGTCTCTCTCATCTCTGTCATCTTGTGATTTTACCATTTGC 61076
Qy 329 PheAlaProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThr 348
Db 61077 TTTGCTCCAAGCAATATTATTCTTATTATTCACATGCTAACTACTACTACTACAACCACT 61136
Qy 349 AspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 368
Db 61137 GATGGCTTATATTTTATATATCTCATAGCTTTGTGCTGGGTAGTCTTATAGTTGCTTA 61196
Qy 369 AspProPheLeuTyrPheValMetSerLys 378
Db 61197 GATCCATTCTTTTATTTCTCATGTCAAAA 61226

RESULT 10
AR070427 AR070427 1102 bp DNA linear PAT 18-FEB-2000
LOCUS DEFINITION Sequence 5 from patent US 5892014.
ACCESSION AR070427

```



```
Qy 149 AsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHi 168
Db 299 AACCTGCCCATCGGGATCTCTTTCTGTGTACAGTGGCGTTTAAAGATCNCCTACCA 358
Qy 168 sLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrThrValValph 188
Db 359 TCTCAATGGCAACAACCTGGGTATTGGCGAGTTCATGTGGCGGATCACCACGGTGGTTT 418
Qy 188 eTyrGlyAsnMetTyrCysAla----IleLeuIleLeuThrCysMetGlyIleAsnArgTy 207
Db 419 CTACGGCAACATGATCTACGCGTANNNTCCCTGATCCTCCTGATGGCATCAACCGCTA 478
Qy 207 rLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLe 227
Db 479 CTTGGCCACGGCTCACCCTTTCACATACCAGAACCTGCCAAGCATGCTTCCATGCT 538
Qy 227 uMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuL 247
Db 539 CATGTGTGGCATGTGTGGGTCTATGTTTCTTATACATGCTGCCCTTGTCAACGNNNA 598
Qy 247 ysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaC 267
Db 599 AGCAGGAGTACACCTCGTCCACTCCGAGATCACCATCTGCCAGCATGCTCGTCAACGCT 658
Qy 267 ysGluSerProSerSerPheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheL 287
Db 659 CGGANTCCCATCATCTCTCCGATTTACTACTTCGTCTCTTACGATTTCTTGGGTTC 718
Qy 287 euIleProPheValIleIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSerL 307
Db 719 TCATCCCGTTGTGATCATCATCTCTCTGTACACGACTCTCATCCACAACTTAAATCAA 778
Qy 307 ysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuIleValIlePheThrI 327
Db 779 AAGATCNGATATGCTGGCTGATACATCAAGCGCGTCTCTCATCTTGTGAATTTACCA 838
Qy 327 leCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyrTyrTyrHisA 347
Db 839 TCTGCTTCCCCCACCACAG----- 858
Qy 347 snThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerC 367
Db 858 ----- 858
Qy 367 ysLeuAspProPheLeuTyrPheValMetSerLysValValAspGlnLeuAsnPro**S 387
Db 858 ----- 858
Qy 387 erAlaMetAlaArgProLeu**ArgProArgArgAspIleTyrGluAspIleHisAlaT 407
Db 859 -----NNNNNNGATATCTGGGAAGACGTACATGCTT 889
Qy 407 rp 407
Db 890 GG 891
RESULT 12
MMPAR2MR
LOCUS M.musculus mRNA for proteinase activated receptor 2.
DEFINITION 2713 bp mRNA linear ROD 08-JUN-1995
ACCESSION Z48043
VERSION Z48043.1 GI:663020
KEYWORDS proteinase activated receptor 2.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2713)
Nystedt, S., Larsson, A.K., Aberg, H. and Sundelin, J.
The mouse proteinase-activated receptor-2 cDNA and gene. Molecular
cloning and functional expression
J. Biol. Chem. 270 (11), 5950-5955 (1995)
MEDLINE 95197620
PUBMED 7890726
REFERENCE 2 (bases 1 to 2713)
AUTHORS Nystedt, S.
TITLE Direct Submission
JOURNAL Submitted (19-JAN-1995) Sverker Nystedt, Division of Molecular
Neurobiology, The Wallenberg Laboratory, Lund University,
Soelvegatan 33A, Lund, S-220 07, Sweden
FEATURES
Location/Qualifiers
1..2713 /organism="Mus musculus"
/db_xref="taxon:10090"
/clone="cm140pBS"
1..72 /citation=[1]
73..1272 /citation=[1]
/codon_start=1
/product="proteinase activated receptor 2"
/protein_id="CAA88097.1"
/db_xref="GI:663021"
/db_xref="MGD:MGI:101910"
/db_xref="SWISS-PROT:P55086"
/translation="MRSLSLAWLGGITLLAASVSCSRSTENLAPGRNNSKGRSLIGRL
ETQPPITKGVPVEPGFSIDEPSASILTKLATVFLPVVYIIVFVIGLPSNGMALWIF
LRTKKKHPAVIYMANLADLSLWIFPKISYHLHGNWVYGEALCKVLIGFYGN
MYCSILFMTCLSVQRYWLVNPKMHPKKAIVGSLAILLIPLVTIPLYVMKQTI
YIPALNITTCDBVLPPEVLVGMFNFSLAIGVFLPALLTASAYVLMIKTLRSSAM
DHSEKRRQRAIRLIITVLANMYFICAPSNLLVHVHYFLIKTORQSHVVALYVALUCL
STLNSCIDPFVYFVSKOPFRHARNALLCRSVRTVNRMQISLSSNFKFSKSGSYSSSS
TSVKTSY"
163..164 /note="genomic intron location"
/citation=[1]
1273..2713 /citation=[1]
2219..2224 /citation=[1]
polyA_signal 2219..2224
polyA_signal 2693..2698
BASE COUNT 632 a 676 c 633 g 772 t
ORIGIN
intron
3'UTR
polyA signal
polyA signal
BASE COUNT 632 a 676 c 633 g 772 t
ORIGIN
Alignment Scores:
Pred. No.: 3..59e-38 Length: 2713
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservative: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 10 Gaps: 10
US-09-208-629f-3 (1-407) x MMPAR2MR (1-2713)
Qy 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuLeuValAlaAlaGlyLeu 27
Db 61 CCACGTCGCGGATGCGAAGTCTCAGCTGGCGTGGCTGCTG-----GGAGGTATC 111
Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
Db 112 ACCCTTCTGGCGCCTCTGGTCTCTCTGACCGGACCGAAGACCTTGACCGGACCGCAAC 171
Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
Db 172 AACAGTAAAGGAAGAAGTCTTATTGGCAGATTAGAAAACCCAGCTCCAATCACTGGGAAA 231
Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGluGly 73
Db 232 GGGGTTCCGGTAGAACGAGCTTTTCCATGATGAGTCTCTCGCTCCATCTCACCGGG 291
Qy 74 TrpThrGlyAlaThrThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
Db 291 ----- 291
Qy 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
Db ----- 113
```



```

Db 790 GACATGTTCAATTACTCTCTCACTGGCCATTGGAGTCTTCTGTTCCCGCCCTCCTT 849
Qy 293 IleIlePheCysTyrThrLeuLeu-----His 302
Db 850 ACTGCATCTGCTAGTCTCATGATCAAGACGCTCGCTCTTCTGTATGGATGAACAC 909
Qy 303 LysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeu 322
Db 910 TCAGAGAGAAAAGCGAGAGGCT-----ATCCGACTCATCATCACCCTGCTG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsn 342
Db 958 GCCATGACTTCAATCTGCTTTCCTAGCAACCTTCTGCTGAGTGCATTATTTCCTA 1017
Qy 343 TyrTyrThrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db 1018 ATCAAAACCCAGAGCGACGCTACGCCCTCTACCTTGTGCGCCCTCTGCTGTCG 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1078 ACCCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGTCTCAAAA 1125

RESULT 14
ARI171259 2732 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 60 from patent US 6297026.
ACCESSION ARI171259
VERSION ARI171259.1 GI:17910209
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE 1 (bases 1 to 2732)
AUTHORS Sundelin,J. and Scarborough,R.M.
TITLE Nucleic acids encoding the C140 receptor
JOURNAL Patent: US 6297026-A 60 02-OCT-2001;
FEATURES
source 1. 2732
/organism="unknown"

BASE COUNT 650 a 676 c 633 g 773 t
ORIGIN

Alignment Scores:
Pred. No.: 3.62e-38 Length: 2732
Score: 584.00 Matches: 128
Percent Similarity: 51.52% Conservative: 76
Best Local Similarity: 32.32% Mismatches: 126
Query Match: 27.34% Indels: 66
DB: 6 Gaps: 10

US-09-208-629F-3 (1-407) x ARI171259 (1-2732)
Qy 8 ProValAlaGlySerGlnAspIleLysMetLysIleLeuValAlaAlaGlyLeu 27
Db 61 CCAGCTCCGGGATCGAAGTCTCAGCTGGCGTGGCTGCTG-----GGAGGTATC 111
Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
Db 112 ACCCTTCTGGCGCTCGGTCTCTCTGAGCCGAGACCTTGACCGGGACGGAAC 171
Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
Db 172 AACAGTAAAGGAGAGAGTCTTATTGGCAGATTAGAAACCCAGCTCCCAATCAGTGGGAA 231
Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluPheProLeuSerAspIleGluGly 73
Db 232 GGGGTTCCGGTAGAACAGGCTTTTCCATCGATGAGTCTCTGCTGCTCATCTCCACCGGG 291
Qy 74 TrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
Db 291 -----
Qy 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113

```

```

Db 292 -----AAGCTGACCACGGTCTTCTTCCG 315
Qy 114 AlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLys 133
Db 316 GTCGTCTACATTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 375
Qy 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
Db 376 TTCTTTTTCGACGAGAGAAACACCCCGCGTGTATTACATGGCCAACTTGGCCCTG 435
Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
Db 436 GCCGACCTCTCTCTGTCATCTGTTCCCTCAAGATCTCTTACCACCTACATGGCAAC 495
Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
Db 496 AACTGGGTCTACGGGAGGCCCTGTGCAAGGTCTCATTTGGCTTTTCTATGTTAAATG 555
Qy 193 TyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
Db 556 TATTGTCCATCTCTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615
Qy 213 ProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Db 616 CCCATGGGACACCCAGG---AAGAAGGCAACATCGCGTGGCGTCTCTTGGCAATC 672
Qy 233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeu 252
Db 673 TGGCTCTGATTTTCTGGTCAACATCCCTTTGTATGTCATGAAGCAGACCATCTACATT 732
Qy 253 ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Db 733 CCAGCATTTGAACATCACCATCTGTACGATGCTGCTGCTGAG---GAGGTATTGGTGGG 789
Qy 273 PheArgPheTyrThrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Db 790 GACATGTTCAATTACTTCTCTCACTGGCCATTGGAGTCTTCTGTTCCCGCCCTCCTT 849
Qy 293 IleIlePheCysTyrThrThrLeuLeu-----His 302
Db 850 ACTGCATCTGCTACGCTCATGATCAAGACGCTCGCTCTTCTGCTATGGATGAACAC 909
Qy 303 LysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeu 322
Db 910 TCAGAGAGAAAAGCGAGAGGCT-----ATCCGACTCATCATCACCCTGCTG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
Db 958 GCCATGACTTCAATCTGCTTTCCTCTAGCAACCTTCTGCTGCTAGTGCATTATTTCCTA 1017
Qy 343 TyrTyrThrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Db 1018 ATCAAAACCCAGAGCGACGCTACGCCCTCTACCTTGTGCGCCCTCTGCTGTCG 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1078 ACCCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGTCTCAAAA 1125

RESULT 15
187850
LOCUS 187850 2732 bp DNA linear PAT 10-AUG-1998
DEFINITION Sequence 60 from patent US 5716789.
ACCESSION 187850
VERSION 187850.1 GI:3407790
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2732)
AUTHORS Sundelin,J. and Scarborough,R.M.
TITLE Method to determine ligands, agonist and antagonist of C140 receptor

```

JOURNAL Patent: US 5716789-A 60 10-FEB-1998;

FEATURES

Location/Qualifiers
1. .2732
/organism="unknown"

BASE COUNT 650 a 676 c 633 g 773 t
ORIGIN

Alignment Scores:

3.62e-38 Length: 2732
Pred. No.: 584.00 Matches: 128
Score: 51.52% Conservative: 76
Percent Similarity: 32.32% Mismatches: 126
Best Local Similarity: 27.34% Indels: 66
Query Match: 6 Gaps: 10
DB:

US-09-208-629F-3 (1-407) x 187850 (1-2732)

Qy 8 ProValAlaGlySerGlnAspIleLeuMetLysIleLeuLeuValAlaAlaGlyLeu 27
Dy 61 CCAGTCCGGGATGCCAAGTCTCAGCTCGCGTGGCTGCTG-----GGAGGTATC 111
Qy 28 LeuPheLeuProValThrVal---CysGlnSerGlyIleAsnValSer-----Asp 43
Dy 112 ACCCTTCTGGCGGCTCGTCTCTCGAGCGGACCGAGAACCTTGACCGGAGCGCAAC 171
Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
Dy 172 AACAGTAAGGAAGAGTCTATTGGCAGATTAGAACCCAGCTCCATCACTCGGAAA 231
Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluPheProLeuSerAspIleGly 73
Dy 232 GGGGTTCCGGTAGAACACGAGGCTTTTCCATCGATGAGTTCTCTGCTCCATCTCACCGG 291
Qy 74 TrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerLeu 93
Dy 291 ----- 291
Qy 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
Dy 292 -----AAGCTGACCAACCGGCTTCTTCTCCG 315
Qy 114 AlaIleTyrIleLeuPheValValGlyValProSerAsnIleValThrLeuTrpLys 133
Dy 316 GTCGTCTACATATTGTGTTGTGATGTTTGCACGTAATGCGCTCGGCTCTGATC 375
Qy 134 LeuSerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
Dy 376 TTCCTTTTCCGAAGAAAGAAACACCCCGCGCTGATTTACATGGCCACCTGGCCTTG 435
Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
Dy 436 GCGACCTCTCTCTGTCATCTGCTTCCCTCGAAGATCTCTTACCACTACATGGCAAC 495
Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMet 192
Dy 496 AACTGGGTCTACGGGAGGCGCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATG 555
Qy 193 TyrCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
Dy 556 TATTGCTCCATCCCTCTCATGACCTGCTCAGCTGCGAGAGGTACTGGGTGATCGTGAAC 615
Qy 213 ProPheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
Dy 616 CCCATGGACACCCAGG---AAGAGGCAACATCGCGTGGCTCTCTTGGCAATC 672
Qy 233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnIleTyrHisLeu 252
Dy 673 TGGCTCCTGATTTTCTGGTCACCATCTCTTGTATGTCATGAAGACGACCATCTACATT 732
Qy 253 ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
Dy 733 CCAGCATTTGAACATCACACACCTGTCAGCATGTGCTGCTGAG---GAGGTATTGGTGGG 789

Qy 273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
Dy 790 GACATGTCAATTACTTCTCTCACTGGCCATTGGAGTCTTCTGTTCGGGCGCTCTT 849
Qy 293 IleIlePheCysTyrThrThrLeuIle-----His 302
Dy 850 ACTGCATCTCGCTACGCTCATGATCAAGACCGCTCGCTCTTCTGTATGGATGAACAC 909
Qy 303 LysLeuLysSerLysAspArgIleTyrLeuGlyTyrIleLysAlaValLeuLeuLeu 322
Dy 910 TCAGAGAGAGAAAGGAGGAGGCT-----ATCGACTCATCATCCGCTGCTG 957
Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleLeuValIleHisAlaAsn 342
Dy 958 GCCATGCTTCACTCTGCTTGTCTAGCAACCTTCTGCTAGTCATATTATTCCTA 1017
Qy 343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGly 362
Dy 1018 ATCAAAACCCAGAGGAGGAGCCAGCTACGCTCTTACCTTGTGCGCTCTGCTGCTG 1077
Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Dy 1078 ACCCTCAACAGCTGCATAGACCCCTTGTCTATTACTTGTCTCAAAA 1125
RESULT 16
BC025432 2772 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, coagulation factor II (thrombin) receptor-like 1,
DEFINITION clone MGC:29183 IMAGE:5006789, mRNA, complete cds.
ACCESSION BC025432
VERSION BC025432.1 GI:19343971
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIA-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Bernberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Teague, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 39 Row: d Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
FEATURES
source 1. .2772
/organism="Mus musculus"
/db_xref="taxon:10090"

/map="FVB/N-3"
 /clone="MGC:29183 IMAGE:5006769"
 /tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month
 old mouse, Taken by biopsy."
 /clone_lib="NCI CGAP Mam2"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 99.1298

CDS

/codon_start=1
 /product="coagulation factor II (thrombin) receptor-like
 1"

/protein_id="AAH25432.1"

/db_xref="GI:19343972"

/db_xref="LocusID:14063"

/translation="MRSLSLAWLLGGITLLAASVSCRTENLAPGRNNSKRSLSIGRL
 ETQPTITGKGVPEPFGSISDEFASILITGKLTITVFPVYIIIVFVIGLPSNGMALWIF
 LFTKKHAPAVIYMANALADLSVIFPLAIYAHHLGNWVYGEALCKVLIGFFYGN
 MYCSILFMTCLSVQRYVIVPMGHPKRNKIANVGSLSAILWILFIIVTIPLYVMKQTI
 YIPALNITCTHDVPELVGDMFNFLSLAIGVFLFPAILTASAVVLMIKTLRSAM
 DEHEKKRQRAIRLIITVLAMYRICFAPSRLLVVHFLIKTORQSHVVALYVALCL
 STLNSCIDPFPYFVSKDFRDHARNALLCRSVTVNRMQISLSNKRFSKSGSYSSSS
 TSVKTSY"

BASE COUNT 671 a 689 c 638 g 774 t

Alignment Scores:

Pred. No.: 4,44e-38 Length: 2772
 Score: 583.00 Matches: 128
 Percent Similarity: 51.26% Conservative: 75
 Best Local Similarity: 32.32% Mismatches: 127
 Query Match: 27.29% Indels: 66
 DB: 10 Gaps: 10

US-09-208-629F-3 (1-407) x BC025432 (1-2772)

Qy 8 ProValAlaGlySerGlnAspIleLeuMetIleLeuValAlaAlaGlyLeu 27
 |||||
 Db 87 CCAGCTCGGGGATCGAAGTCTCAGCTGCGGTGCTG-----GGAGGTATC 137
 Qy 28 LeuPheLeuProValThrVal-----CysGlnSerGlyIleAsnValSer-----Asp 43
 |||||
 Db 138 ACCCTTCTGCGCGGCTCGGTCTCTCGAGCGGACCGAGAACCTTGACCGGGACGCAAC 197
 Qy 44 AsnSer-----AlaLysProThrLeuThrIleLys 53
 |||||
 Db 198 AACAGTAAGGAAGAAGTCTATTGGCAGATTGGAACCCAGCTCCATCCTCGGAAA 257
 Qy 54 SerPheAsnGlyGlyProGlnAsnThrPheGluPheProLeuSerAspIleGly 73
 |||||
 Db 258 GGGGTTCGGTAGAACACGAGCTTTCCATCGATGAGTCTCTCGCTCCATCTCACCGGG 317
 Qy 74 TrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerThrLeu 93
 |||||
 Db 317 ----- 317
 Qy 94 HisValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIlePro 113
 |||||
 Db 318 -----AAGTGAACACCTGCTTTCTTC 341
 Qy 114 AlaIleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLys 133
 |||||
 Db 342 GTCGTCTACATATTGTTGTGATGTTTGGCCAGTAATGCATGGCCCTCTGGATC 401
 Qy 134 LeuSerLeuArgThr-----LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIle 152
 |||||
 Db 402 TTCCTTTTCCGAACGAAGAAACACCCCGCGGTGATTTACATGGCCAACTCGGCCTTG 461
 Qy 153 AlaAspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 172
 |||||
 Db 462 GCTGACCTCTCTCTGTCATCTGGTTCCCTCGCCATGCTACCACTACATGGGAAC 521
 Qy 173 AsnTrpValPheGlyGluValMetCysArgIleThrThrValPheTyrGlyAsnMet 192
 |||||

Db 522 AACTGGGTCTATGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGCAACATG 581
 Qy 193 TyrCysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHis 212
 |||||
 Db 582 TATTGCTCATCTCTTTCATGACCTCCCTCAGGTGCAGAGTACTGGGTGATCGTGAAC 641
 Qy 213 PropheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleVal 232
 |||||
 Db 642 CCCATGGGACCCACAG--AAGAAGGCAACATCGCGGTGGCTCTCTTGGCAATC 698
 Qy 233 TrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGlyThrHisLeu 252
 |||||
 Db 699 TGGCTCCTGATTTTCTGTCACCATCTCTTGTATGTCATGAAGCAGACCATCTACAT 758
 Qy 253 ValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSer 272
 |||||
 Db 759 CCAGCTTGAACATCACCACTGCCAGATGCTGCCTGAG---GAGGTATTGGTGGGG 815
 Qy 273 PheArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 292
 |||||
 Db 816 GACATGTTCAATTACTTCTCTCAGTGGCATTGGAGTCTTCTGTTCCGGCCATCCT 875
 Qy 293 IleIlePheCysTyrThrThrLeuIle-----His 302
 |||||
 Db 876 ACTGATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTTCTGCTATGATGAAC 935
 Qy 303 LysLeuLysSerLysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuLeu 322
 |||||
 Db 936 TCAGAGAAGAAAAGCAGAGGCT-----ATCGATCATCATCAGCGGTGCTG 983
 Qy 323 ValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsn 342
 |||||
 Db 984 GCCATGTAATCTCATCTGCTTGTCTCTAGCACTTCTGCTGATGCTATTTCCTA 1043
 Qy 343 TyrTyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleLeuAlaLeuCysLeuGly 362
 |||||
 Db 1044 ATCAAAACCCAGAGGAGGAGCCAGCTACGCGCTCTACCTGCTCGCCCTCTGCTGTCG 1103
 Qy 363 SerLeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
 |||||
 Db 1104 ACTCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGCTCTCAAAA 1151
 RESULT 17
 XLU09632 1680 bp mRNA linear VRT 28-MAY-1994
 LOCUS Xenopus laevis thrombin receptor mRNA, complete cds.
 DEFINITION U09632
 ACCESSION U09632.1 GI:495197
 VERSION
 KEYWORDS
 SOURCE African clawed frog.
 ORGANISM
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 420)
 AUTHORS Gersten,R.E., Chen,J., Ishii,M., Ishii,K., Wang,L., Nanevicz,T.,
 Turck,C.W., Vu,T.K. and Coughlin,S.R.
 TITLE Specificity of the thrombin receptor for agonist peptide is defined
 by its extracellular surface
 JOURNAL Nature 368 (6472), 648-651 (1994)
 MEDLINE 94195429
 PUBMED 8145852
 REFERENCE 2 (bases 1 to 1680)
 AUTHORS Coughlin,S.R.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-1994) Shaun R. Coughlin, CVRI, University of
 California, San Francisco, 505 Parnassus Avenue, Room HSW 831, San
 Francisco, CA 94143, USA
 FEATURES
 source Location/Qualifiers
 1..1680 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 197..1459

```

/codon_start=1
/product="thrombin receptor"
/protein_id="AAA18498.1"
/db_xref="GI:495198"
/translation="MMELRVLLLLLLTLLGMSGLCLANSIDTOAGAHNNMTIKTF
RIPDSESEFEIIPWDELDESGSGDOAPVSRSAKPIRNITKEAQVLSQWLTK
FVSLYTVFIVGLPLNLLAIILFLPMKVRPAVYMLNLAIDVFFSVLPLFKIAY
HLSGNLWLFPGMCRIVTAIFYNMYCVSVLLIASISVDRELVAVYPMHSWETMERA
YMACSFILISASTIPLVTEQTKIPRLDITCHDVLKDKLDFYIYFFSFCLL
FFVPFIIITICVIGIIRSLSSSIENSCKKTRALLFLAVVLVCVFIICFQTNVLFIT
HYLQENEFYFAVILSACVGSVCCLDPLIYIYASSQCCQRYLYLSLLCCRKVEPSSS
TGQLMSPAMNDKNSINAKSIVYKLLA"
BASE COUNT 427 a 366 c 401 g 486 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 2,78e-38 Length: 1680
Score: 582.50 Matches: 124
Percent Similarity: 54.35% Conservativeness: 82
Best Local Similarity: 32.72% Mismatches: 138
Query Match: 27.27% Indels: 35
DB: 5 Gaps: 11

```

US-09-208-629F-3 (1-407) x XLU09632 (1-1680)

```

QY 19 IleuIleuLeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGly 38
DB 215 CTGCTGCTGCTGCTGCTATTGACCCCTGCGGAGCGATGGATCTCTCTGCTGGCCAAC 274
QY 39 IleAsnValSerAspAsnSerAlaLys-----ProThrLeuThrIleLysSer 54
DB 275 -----TCTGATACCCAGCAAGGGGTGCTCATTTCAACAATATGACCAATAAAAT 325
QY 55 PheAsn-----GlyGlyProGlnAsnThrPheGluGluPheProLeuSerAspIleGlu 72
DB 326 TTGAGATATTTGACGACAGTGAAGTGGTGTGAAGAGATTCCATCGGATGAATCGAT 385
QY 73 GlyTrpThrGlyAlaThrThrIleLysAlaGluCysProGluAspSerIleSerThr 92
DB 386 GAA---TCAGGAGAAGTTTCAGGAGACCGCCGGTGTCCCGTAGTGCAAGGAGCCCA 442
QY 93 Leu-----HisValAsnAsnAlaThrIleGlyTyrlleuArgSerSerLeuSerThrGln 110
DB 443 ATCCGGAGGAACATCACAAAGGAGCGGAGCAATATCTCTCCAGCTAGTGGCTGACCA 502
QY 111 ValIleProAlaIleTyrlleLeuLeuPheValValGlyValProSerAsn----- 127
DB 503 TTGTGATACCTCCCTGACATGCTGCTGTTATGTGGGCTGCTCTGAATCTGTGGCA 562
QY 128 ---IleValThrLeuTrpLysLeuSerLeuArgThrLysSerIleSerLeuValIlePhe 146
DB 563 ATCATCATATTCCTGTTCAAGATGAAAGTCAGAAAACCGGCA-----GTGGTCTAT 613
QY 147 HisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLysIleAla 166
DB 614 ATGTGAACCTGGCCATCGCCGACGCTGTTCTTTAGTGTGTTGCTTTTCAAAATAGCC 673
QY 167 TyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThrVal 186
DB 674 TACCATCTGTCGGAACGACTGCTGTTTGGGCTGGATGTGCCCATGTGCACCTGCC 733
QY 187 ValPheTyrlleGlyAsnMetTyrlleCysAlaIleLeuIleLeuThrCysMetGlyIleAsnArg 206
DB 734 ATCTTCTACTGCAACATGTAAGTCTGCTGCTGCTGCTCATCGCAGCATCAGTGTGGACAGG 793
QY 207 TyrlleuAlaThrAlaHisProPheThrTyrlleGlnLysLeuProLysArgSerPheSerLeu 226
DB 794 TTCTGGCAGTGGTGTATATCCCATCACTCCCTTTCTTGGGCGCAATATGAGCCGCTGCTAT 853
QY 227 LeuMetCysGlyIleValTrpValMetValPheLeuTyrlleMetLeuProPheValIleLeu 246
DB 854 ATGGCCTGTGATTTTCATTTGCTGATATCCATGAGGAGCATATATACCATCTTCTGTGACT 913

```

```

QY 247 LysGlnGluTyrlleHisLeuValHisSerGluIleThrThrCysHisAspValValAspAla 266
DB 914 GAGCAAACTCAGAAAGATTCTAGGCTGGATATTACTACTTGGCATGATGTTTAGAT--- 970
QY 267 CysGluSerProSerSerPheArgPheTyrllePheValSerLeuAlaPhePheGlyPhe 286
DB 971 TTGAAAGACTCMAAGACTTTTATATCTACTATTCTTCTTCTTCTTCTGCTGCTGCTTCT 1030
QY 287 LeuIleProPheValIleIleIlePheCysTyrlleThrThrLeuIleHisLysLeuLysSer 306
DB 1031 TTGCTGCCATTCAATATCACCACATCTCTGATACAGGATCATCAGAAGCTTAAGTTC 1090
QY 307 -----LysAspArgIleTrpLeuGlyTyrlleLysAlaVal 318
DB 1091 TCCAGCATCGAAGCAAGTTGCAAGAAAGACAGACC-----CTGTCTTCTGGCG 1138
QY 319 LeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIleIleLeuValIle 338
DB 1139 GTGGTTGCTGCTGCTGCTTTTATCATCTGTTTGGGCCAACCAATGCTCTTCTTCTGACC 1198
QY 339 HisHisAlaAsnTyrlleTyrlleHisAsnThrAspSerLeuTyrlleMetTyrlleAla 358
DB 1199 CAT-----TACTTGAAGAGCAATGAATTTCTATATTTGCTTACATTCGAGT 1249
QY 359 LeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrlleValMetSer 377
DB 1250 GCCTGTGTTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1306
RESULT 18
MMPROACR2 1477 bp DNA linear ROD 30-JAN-1995
LOCUS M.musculus exon 2 for proteinase activated receptor 2.
ACCESSION 235158
VERSION 235158.1 GI:559918
KEYWORDS proteinase activated receptor; proteinase activated receptor 2.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1477)
AUTHORS Nystedt,S., Emilsson,K., Wahlestedt,C. and Sundelin,J.
TITLE Molecular cloning of a potential proteinase activated receptor
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (20), 9208-9212 (1994)
MEDLINE 95023880
PUBMED 7937743
REFERENCE 2 (bases 1 to 1477)
AUTHORS Nystedt,S.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-1994) Sverker Nystedt, Division of Neurobiology,
The Wallenberg, Laboratory, Lund University, Soelvegatan 33A, Lund,
S-220 07, Sweden
REFERENCE 3 (bases 1 to 1477)
AUTHORS Nystedt,S., Larsson,A.K., Aaberg,H. and Sundelin,J.
TITLE The mouse proteinase activated receptor 2 cDNA and gene-molecular
cloning and functional expression
JOURNAL J. Biol. Chem. 170 (1995) In press
FEATURES
Location/Qualifiers
source 1..1477
/organism="Mus musculus"
/db_xref="taxon:10090"
/germline
gene 1..1477
/gene="PAR-2"
intron <1..310
/gene="PAR-2"
/citation=[3]
exon 311..1477
/gene="PAR-2"
/citation=[3]
/number=1
3'UTR 1420..1477
/gene="PAR-2"
BASE COUNT 341 a 413 c 321 g 402 t

```


Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 174361 CTTTTCGGAACGAAGAAGAACCCCGCGCTGATTTACATGGCCAACTGGCTGGCC 174302
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyHisLeuAsnGlyAsnAsn 173
Db 174301 GACCTCTCTCTGTCTATCTGGTCCCTCGAGATCTCCTACCACTTACATGGCAACAAC 174242
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyHisAsnMetTyr 193
Db 174241 TGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATCTGGCTTTTCTATGGTAACATGTAT 174182
Qy 194 CysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyHisLeuAlaThrAlaHisPro 213
Db 174181 TGTCTCATCTCTTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCC 174122
Qy 214 PheThrTyHisLeuLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
Db 174121 ATGGGACACCCCGAGG---AAGAAGGCAACATCCCGCTGGCGTCTCTTGGCAATCTGG 174065
Qy 234 ValMetValPheLeuTyMetLeuProPheValIleLeuLysGlnTyHisLeuVal 253
Db 174064 CTCTGATTTTCTGGTTCACCATCTCTTGTATGTCATGAGCAGACCATCTACATTTCCA 174005
Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
Db 174004 GCATTGAACATCACCATCTGTACCATGTGCTGCTGAG---GAGGTATTGGTGGGGAC 173948
Qy 274 ArgPheTyHisLeuValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
Db 173947 ATGTTCATTAATTTCT 173888
Qy 294 IlePheCysTyHisThrThrLeuIle-----HisLys 303
Db 173887 GCATCTCGCTACGGTCTCATGATCAAGCGCTCGCTCTCTCTCTCTCTCTCTCTCTCT 173828
Qy 304 LeuLysSerIleAspArgIleTrpLeuGlyTyHisLysAlaValLeuLeuVal 323
Db 173827 GAGAAGAAAGGAGAGAGGCT-----ATCGACATCATCATCATCATCATCATCATCAT 173780
Qy 324 IlePheThrIleCysPheAlaProThrAsnIleIleValIleHisLeuAlaAsnTyr 343
Db 173779 ATGTTCATTAATTTCT 173720
Qy 344 TyHisAsnThrAspSerLeuTyHisPheMetTyHisLeuAlaLeuCysLeuGlySer 363
Db 173719 AAAACCCAGAGGAG 173660
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyHisPheValMetSerTyr 378
Db 173659 CTCACAGCTGCATAGACCCCTTTGCTATTACTTTGCTCTCAAAA 173615

RESULT 20

AR012637 LOCUS 1475 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 1 from patent US 5763575.
ACCESSION AR012637
VERSION AR012637.1 GI:3970627
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1475)
AUTHORS Sundelin, J. and Scarborough, R. M.
TITLE Agonist and antagonist peptides of the C140 receptor
JOURNAL Patent: US 5763575-A 1 09-JUN-1998;
FEATURES Location/Qualifiers
source 1. 1475
/organism="unknown"

BASE COUNT 341 a 414 c 319 g 401 t
ORIGIN

Alignment Scores:

Pred. No.: 4,21e-38 Length: 1475
Score: 579.50 Matches: 111
Percent Similarity: 60.00% Conservative: 66
Best Local Similarity: 37.63% Mismatches: 101
Query Match: 27.13% Indels: 17
DB: Gaps: 5
US-09-208-629F-3 (1-407) x AR012637 (1-1475)
Qy 95 ValAsnAsnAlaThrIleGlyTyHisLeuSerSerLeuSerThrGlnValIleProAla 114
Db 406 ATCGATGAGTCTTCGGTCCATCTCACCGGAAGTCACCGGTCTTCTTCCTCCGGTC 465
Qy 115 IleTyHisLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
Db 466 GTCTACATTATTGTTTGTGATTGTTTGCCTCAGTAATGGCATGGCCCTCTGGATCTTC 525
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 526 CTTTTCGGAACGAAGAAGAAACACCCCGCGTGATTATACATGGCCACCTGGCCCTGGCC 585
Qy 154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyHisLeuAsnGlyAsnAsn 173
Db 586 GACCT 645
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyHisLeuValTyr 193
Db 646 TGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATCTGGCTTTTCTATGGTAACATGTAT 705
Qy 194 CysAlaIleLeuLeuLeuThrCysMetGlyIleAsnArgTyHisLeuAlaThrAlaHisPro 213
Db 706 TGTCTCATCTCTTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCC 765
Qy 214 PheThrTyHisLeuLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
Db 766 ATGGGACACCCCGAGG---AAGAAGGCAACATCCCGCTGGCGTCTCTTGGCAATCTGG 822
Qy 234 ValMetValPheLeuTyMetLeuProPheValIleLeuLysGlnTyHisLeuVal 253
Db 823 CTCTGATTTTCTGGTCAACCATCTCTTGTATGTCATGAAGCAGACCATCTACATTTCCA 882
Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
Db 883 GCATTGAACATCACCATCTGTACGATGTCTGCTGAG---GAGGTATTGGTGGGGAC 939
Qy 274 ArgPheTyHisLeuValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
Db 940 ATGTTCAATTTACTTCT 999
Qy 294 IlePheCysTyHisThrThrLeuIleHisLysLeuLysSer----- 306
Db 1000 GCATCTGCCTACGTGCTCATGATCAAGACGCTCGCTCTCTCTCTCTCTCTCTCTCTCTCT 1059
Qy 307 -----LysAspArgIleTrpLeuGlyTyHisLysAlaValLeuLeuVal 323
Db 1060 GAGAACAAGGAGGAGAGGCT-----ATCGACATCATCATCATCATCATCATCATCATCAT 1107
Qy 324 IlePheThrIleCysPheAlaProThrAsnIleIleValIleHisLeuAlaAsnTyr 343
Db 1108 ATGATTTCTATCTGCTTTGCTCTAGCAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1167
Qy 344 TyTyHisAsnThrAspSerLeuTyHisPheMetTyHisLeuAlaLeuCysLeuGlySer 363
Db 1168 AAAACCCAGAGGAG 1227
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyHisPheValMetSerTyr 378
Db 1228 CTCACAGCTGCATAGACCCCTTTGCTATTACTTTGCTCTCAAAA 1272

RESULT 21

AR171257 LOCUS 1475 bp DNA linear PAT 17-DEC-2001

DEFINITION Sequence 1 from patent US 6297026.
ACCESSION AR171257
VERSION AR171257.1 GI:17910207
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1475)
AUTHORS Sundelin,J. and Scarborough,R.M.
TITLE Nucleic acids encoding the C140 receptor
JOURNAL Patent: US 6297026-A 1 02-OCT-2001;
FEATURES Location/Qualifiers
source 1..1475
BASE COUNT 341 a 414 c 319 g 401 t
ORIGIN

Alignment Scores:
Pred. No.: 4,21e-38 Length: 1475
Score: 579.50 Matches: 111
Percent Similarity: 60.00% Conservative: 66
Best Local Similarity: 37.63% Mismatches: 101
Query Match: 27.13% Indels: 17
DB: 6 Gaps: 5

US-09-208-629F-3 (1-407) x AR171257 (1-1475)

Qy 95 ValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db 406 ATCGATGAGTTCTCGTCCATCTCCACCGGAAGCTGACCGAGTCTTCTTCCGGTC 465
Qy 115 IleTyrIleLeuLeuPheValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
Db 466 GTCTACATTATTGTGTTTGTGATTGTTGGCCAGTAATGGCATGGCCCTCTGGATCTTC 525
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 526 CTTTTCGACGAGGAAGAAACACCCCGCGTGATTACATGGCCACCTTGGCCCTTGGCC 585
Qy 154 AspleuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 586 GACCTCTCTCTGTCTATCTGTGTTCCCTCCAGTAATGGCATGGCCCTCTGGATCTTC 645
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db 646 TGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTTCATGTAACATGAT 705
Qy 194 CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGCTCCATCTCTCTACATGACCTGCTCAGCGTCAGAGGTACTGGGTGATCGTGAACCCC 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
Db 766 ATGGGACACCCCGAG---AAGAAAGCAACATCGCCGTGGCGTCTCTTGGCAATCTGG 822
Qy 234 ValMetValPheLeuTyrMetLeuPheValIleLeuLysGlnGluTyrHisLeuVal 253
Db 823 CTCTGATTTTCTGGTCACCATCTTGTATGTCATGAGAGACCACTACATTTCCA 882
Qy 254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
Db 883 GCATTGAACATCACCATCTGTCAGATGTGTCGCTGAG---GAGGTATTGGTGGGGAC 939
Qy 274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIle 293
Db 940 ATGTTCATTAATCTCTCTACTGCGCATTTGGAGTCTTCTCTGTTCCCGGCCCTCTTACT 999
Qy 294 IlePheCysTyrThrThrLeuIleHisLysLysSer----- 306
Db 1000 GCATCTGCCTACGTCTCATGATCAAGACGCTCCGCTCTTCTGCTATGGATGAACACTCA 1059
Qy 307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323

1060 GAGAACAAAAGGCAGAGGGCT-----ATCCGACTCATCATCCGCTGTGGCC 1107
Qy 324 IlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyr 343
Db 1108 ATGTACTTCTCTCTCTAGCAACCTTCTCTAGTAGTCATTATTTCCCTAATC 1167
Qy 344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySer 363
Db 1168 AAAACCCAGAGGCAGGCCAGCTTACGCCCTTACCTTGTGCGCTCTGCTGTGCGACC 1227
Qy 364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db 1228 CTCAACAGCTGCATAGACCCCTTGTCTATTACTTGTCTCAAAA 1272

RESULT 22
142454
LOCUS
DEFINITION Sequence 1 from patent US 5629174.
ACCESSION 142454
VERSION 142454.1 GI:2467949
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1475)
AUTHORS Sundelin,J. and Scarborough,R.M.
TITLE Recombinant C140 receptor
JOURNAL Patent: US 5629174-A 1 13-MAY-1997;
FEATURES Location/Qualifiers
source 1..1475
BASE COUNT 341 a 414 c 319 g 401 t
ORIGIN

Alignment Scores:
Pred. No.: 4,21e-38 Length: 1475
Score: 579.50 Matches: 111
Percent Similarity: 60.00% Conservative: 66
Best Local Similarity: 37.63% Mismatches: 101
Query Match: 27.13% Indels: 17
DB: 6 Gaps: 5

US-09-208-629F-3 (1-407) x 142454 (1-1475)

Qy 95 ValAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db 406 ATCGATGAGTTCTCGTCCATCTCCACCGGAAGCTGACCGAGTCTTCTTCCGGTC 465
Qy 115 IleTyrIleLeuLeuPheValGlyValProSerAsnIleValThrLeuTyrLysLeu 134
Db 466 GTCTACATTATTGTGTTTGTGATTGTTGGCCAGTAATGGCATGGCCCTCTGGATCTTC 525
Qy 135 SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db 526 CTTTTCGACGAGGAAGAAACACCCCGCGTGATTACATGGCCACCTTGGCCCTTGGCC 585
Qy 154 AspleuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db 586 GACCTCTCTCTGTCTATCTGTGTTCCCTCCAGTAATGGCATGGCCCTCTGGATCTTC 645
Qy 174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db 646 TGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTTCATGTAACATGAT 705
Qy 194 CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db 706 TGCTCCATCTCTCTACATGACCTGCTCAGCGTCAGAGGTACTGGGTGATCGTGAACCCC 765
Qy 214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTyr 233
Db 766 ATGGGACACCCCGAG---AAGAAAGCAACATCGCCGTGGCGTCTCTTGGCAATCTGG 822
Qy 234 ValMetValPheLeuTyrMetLeuPheValIleLeuLysGlnGluTyrHisLeuVal 253

```
Db      823 CTCCTGATTTTCTGGTCAACATCCCTTTGTATGTCATGAAGCAGCAACATCTACATTTCCA 882
Qy      254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSerPhe 273
Db      883 GCATTGAACATCACCACTGTCACGATGTCGCTGAG---GAGGTATTGGTGGGGGAC 939
Qy      274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
Db      940 ATGTTCAATTACTTCTCTCACTGGCCATTGGAGTCTTCTGTTCCCGGCCCTCTTACT 999
Qy      294 IlePheCysTyrThrThrLeuIleHisLeuLysSer----- 306
Db      1000 GCATCTGCCGTACGCTCATGATCAAGACGCTCGCTCTTCTGCTATGGATGAACACTCA 1059
Qy      307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323
Db      1060 GAGBACAAAGGCAGAGGCT-----ATCGACTCATCATCAGCGTGGTGGCC 1107
Qy      324 IlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyr 343
Db      1108 ATGACTTCTATCTGCTTGTCTCAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATC 1167
Qy      344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleLeuAlaLeuCysLeuGlySer 363
Db      1168 AAAACCCAGGCGAGCCAGCTACGCTACGCCCTCTACCTTGTGCGCCTCTGCTGTCGACC 1227
Qy      364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db      1228 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 1272

RESULT 23
LOCUS      I87848      1475 bp      DNA      linear      PAT 10-AUG-1998
DEFINITION Sequence 1 from patent US 5716789.
ACCESSION  I87848
VERSION    I87848.1 GI:3407788
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 1475)
AUTHORS    Sundelin,J. and Scarborough,R.M.
TITLE      Method to determine ligands, agonist and antagonist of C140
           receptor
JOURNAL    Patent: US 5716789-A 1 10-FEB-1998;
FEATURES   Location/Qualifiers
           1..1475
           /organism="unknown"
BASE COUNT 341 a 414 c 319 g 401 t
ORIGIN

Alignment Scores:
Pred. No.: 4.21e-38      Length: 1475
Score: 579.50      Matches: 111
Percent Similarity: 60.00%      Conservative: 66
Best Local Similarity: 37.63%      Mismatches: 101
Query Match: 27.13%      Indels: 17
DB: 6      Gaps: 5

US-09-208-629F-3 (1-407) x I87848 (1-1475)

Qy      95 ValAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla 114
Db      406 ATCGATGAGTCTCTGGGTCCATCTCACCGGAGCTGACCGGCTTTCTTCCGGTC 465
Qy      115 IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpLysLeu 134
Db      466 GTCTACATTATTGTGTTGTGATTGTTGCCCCAGTAATGCGATGCCCTCTGATCTTC 525
Qy      135 SerLeuArgThr-----LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla 153
Db      526 CTTTTCGACGAAGAAGAAACACCCCGCGGTGATTTTACATGGCCAACTGGCCTTGGCC 585
```

```
Qy      154 AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn 173
Db      586 GACCTCTCTCTCTGTCATCTGGTTCCCTCGAGAGATCTCTTACCACCTACATGCAACAAC 645
Qy      174 TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr 193
Db      646 TGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATGTAT 705
Qy      194 CysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro 213
Db      706 TGTCTCATCTCTTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCC 765
Qy      214 PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrp 233
Db      766 ATGGGACACCCACAG---AAGAAGGCAACATCGCGTGGCTCTCTTGGCAATCTGG 822
Qy      234 ValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyrHisLeuVal 253
Db      823 CTCCTGATTTTCTGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATTCCA 882
Qy      254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerPhe 273
Db      883 GCATTGAACATCACCACTGTCACGATGCTCCTGAG---GAGGTATTGGTGGGGGAC 939
Qy      274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIle 293
Db      940 ATGTTCAATTACTTCTCTCAGTGGCATTGGAGTCTTCTGTTCCCGGCCCTCTTACT 999
Qy      294 IlePheCysTyrThrThrLeuIleHisLysLeuLysSer----- 306
Db      1000 GCATCTGCTACGTCGTCATGATCAAGACGCTCGCTCTTCTGCTATGGATGAACACTCA 1059
Qy      307 -----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuVal 323
Db      1060 GAGAACAAAGGCAGAGGCT-----ATCGACTCATCATCAGCGTGGTGGCC 1107
Qy      324 IlePheThrIleCysPheAlaProThrAsnIleIleLeuValIleHisAlaAsnTyr 343
Db      1108 ATGACTTCTATCTGCTTGTCTCTAGCAACCTTCTGCTCGTAGTGCATTATTTCCTAATC 1167
Qy      344 TyrTyrHisAsnThrAspSerLeuTyrPheMetTyrIleLeuAlaLeuCysLeuGlySer 363
Db      1168 AAAACCCAGGCGAGCCAGCTACGCTACGCCCTCTACCTTGTGCGCCTCTGCTGTCGACC 1227
Qy      364 LeuAsnSerCysLeuAspProPheLeuTyrPheValMetSerLys 378
Db      1228 CTCACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAA 1272
```

```
RESULT 24
LOCUS      BC012453      2813 bp      mRNA      linear      PRI 20-AUG-2001
DEFINITION Homo sapiens, clone MGC:9549 IMAGE:3857382, mRNA, complete cds.
ACCESSION  BC012453
VERSION    BC012453.1 GI:15214649
KEYWORDS   MGC.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens.
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 2813)
AUTHORS    Strausberg,R.
TITLE      Direct Submission
JOURNAL    Submitted (15-AUG-2001) National Institutes of Health, Mammalian
           Gene Collection (MGC), Cancer Genomics Office, National Cancer
           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
           USA
REMARK     NIH-MGC Project URL: http://mgc.nci.nih.gov
           Contact: MGC help desk
           Email: cgabbs-f@mail.nih.gov
           Tissue Procurement: DCTD/DTP
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalon@bcm.tmc.edu
 Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 21 Row: b Column: 20
 This clone was selected for full length sequencing because it passed the following selection criteria: GenomesCan gene prediction, Similarity but not identity to protein.

FEATURES

source

1. 2813

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:9549 IMAGE:3857382"

/tissue_type="Ovary, adenocarcinoma"

/clone_lib="NIH_MGC_66"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

93. 1286

/codon_start=1

/product="Unknown (protein for MGC:9549)"

/protein_id="AAH12433.1"

/db_xref="GI:15214650"

/translation="MRSPSAALLGAAILLAAALSCSGTQGTNRSGKSLRGKVDG
 TSHVGTGVTETVFSVDEFSASVLTKLTFPIVTVTVFVGVLPNSGMALWFLF
 RTKKHPAVIYMLADLLSVLWPLKJAYHIGNNVIYGEALCNVLIGFPYGNMY
 CSLLFMTCISVQVWVTVNPKHSHKKNATIGISLAILLILVPLVVKOTIFI
 PALNITTCHEVLPEQLLVGMFNFLSLGCVLPFAPFLASVYLMVIMRLRSANDE
 NSEKRRRAIKLIVTLVMIYLICTPSNLLLVHFLIKSGQSHVYLVIVLALCLST
 LNSCIDPFVYFVSHDFRDHAKNALLCRSVRTKMQVSLTSKKHRSKSSSSSTT
 VKTSY"

BASE COUNT 757 a 614 c 604 g 838 t

ORIGIN

Alignment Scores:

Prod. No.: 6.17e-37 Length: 2813
 Score: 569.00 Matches: 120
 Percent Similarity: 52.59% Conservative: 73
 Best Local Similarity: 32.70% Mismatches: 136
 Query Match: 26.64% Indels: 38
 Db: 9 Gaps: 8

US-09-208-629F-3 (1-407) x BC012453 (1-2813)

Qy 22 LeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
 Db 117 CTGCTGGGGCGCCATCTCTAGACGCTCTCTCTCTGACGTGGACCATCAAGGA 176
 Qy 42 SerAsnSerAlaGlyProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
 Db 177 ACCATAGATCTCTTAAGGAGAGCGCTTATGTAAGTTGATGCG----- 224
 Qy 62 ThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIle 81
 Db 225 -----ACATCCACGTC-----ACTGAAAGAGGATGTACAGTT 257
 Qy 82 LysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGly 101
 Db 258 GAA-----ACAGTCTTTCTGGGATGAGTTTCTGCACT 293
 Qy 102 TyrLeuArgSerSerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 121
 Db 294 GTCCTCACTGGAACACTGACACTGCTCTCTCTCAATGCTACACAAATGTGTTGTG 353
 Qy 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSer 140
 Db 354 GTGGGTTTCCCAAGTAACGGCATGCCCTGTGGGTCTTCTTTTCCGAACTAAGAGAA 413

Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
 Db 414 CACCCTGCTGTGATTACATGCGCAATCTGGCTTGGCTGACCTCTCTCTGTCATCTGG 473
 Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
 Db 474 TTCCCTTGAAGATTGCTTATCACATACATGCGCAACACTGGATTTATGGGAAGCTCTT 533
 Qy 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThr 200
 Db 534 TGAATGTGCTTATTGGCTTTTCTATGGCAACATGACTGTCTTCATTTCTTCATGACC 593
 Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrTrpGlnLysLeuPro 220
 Db 594 TGCCTCAGTGTGCAGAGGTATTGGGTATCGTGAACCCCATGGGGCACCACCG---AAG 650
 Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
 Db 651 AAGGCAACATTCGCCATTCCTCGGCAATATGCTGCTGATTTCTGCTGGTCACC 710
 Qy 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
 Db 711 ATCCCTTGTATGTCGTGAAGCAGACCATCTTCAATTCCTGCCCTGAACATCACGACCTGT 770
 Qy 261 HisAspValValAspAlaCysGluSerProSerPhe-----ArgPheTyr 276
 Db 771 CATGATGTTTG-----CCTGAGCAGCTCTTGTGGGAGACATGTTCAAT 815
 Qy 277 TyrPheValSerLeuAlaPhePheGlyLeuLeuPheValIleIlePheCys 296
 Db 816 TACTTCTCTCTGCGCATTCGGGTCTTTCTGTTCCAGCCCTCTCTCACAGCTCTGCC 875
 Qy 297 TyrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
 Db 876 TATGCTGCTGATGATGAGATGTCGCAATCTTCCCATGATGATAAACTCAGAGAAGAA 935
 Qy 311 TrpLeuGlyTyrIleLysAlaValLeuLeuValIlePheThrIleCysPheAla 330
 Db 936 AGGAAGAGGCGCCATCAACTCATTTGCTGCTGCTGCGCATGATGATGATGCTGCTCACT 995
 Qy 331 ProThrAsnIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
 Db 996 CTTAGTAACCTCTGCTTGTGGTGCATTTATTTCTGATTAAAGACCGAGGCCAGAGCCAT 1055
 Qy 351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
 Db 1056 GTCTATGCCCTGTACATTTAGCCCTCTGCTCTCTACCTTAAACAGCTGATCGACCCC 1115
 Qy 371 PheLeuTyrPheValMetSer 377
 Db 1116 TTGTCTATTACTTTGTTCA 1136

RESULT 25

LOCUS

DEFINITION

BC018130 2876 bp mRNA linear PRI 06-DEC-2001
 Homo sapiens, coagulation factor II (thrombin) receptor-like 1,
 clone MGC:9298 IMAGE:3895653, mRNA, complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REMARK

REMARK

REMARK

REMARK

REMARK

REMARK

REMARK

REMARK

REMARK

REMARK

REMARK

JOURNAL Br. J. Pharmacol. 118 (3), 521-530 (1996)
MEDLINE 96358009
PUBMED 8762073

REFERENCE 2 (bases 1 to 1428)

AUTHORS Al-Ani,B., Hollenberg,M.D., Saifeddine,M., Cheng,C.-H. and Wang,L.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1996) Pharmacology, Univ. of Calgary, 3330 Hospital Dr NW, Calgary, AB T2N 4N1, Canada

FEATURES Location/Qualifiers
source I..1428
organism="Rattus norvegicus"
strain="Sprague-Dawley"
db_xref="taxon:10116"
tissue_type="intestinal and kidney"

CDS I..1194
note="PAR-2"
codon_start=1
product="proteinase-activated receptor-2"
protein_id="AAC52703.1"
db_xref="GI:A480688"
translation="MRSLSLAWLGGITLLAASACNRVNPAPGNSKGRSLIGRLDT
PPPTTKGAPVPGSVDFSAVLTVFKUTTFPLFYIIIVFGLPSNGMALWVFFF
RTKKHGAVYMANALADLLSWIFPEKISYHLHGNDWTGDALCKVLIGFYFGNIY
CSILFTWCQSQRVYIWNPMGHRSKRANIAVGSLAIWLFIPLVTIPLYVMROGTIY
PALNTTCHDVLPEEVLDGMFSYFLSLAIGVFLFPALLTAGAYLMIKTLRSSAMDE
HSEKKERRAIRLIITVLSMYFICFAPSNNLVAVHFELIKSQSHVYALYLVALCLST
LNKCIDPFVYYFVKDPDRQARNALLCRSVRTVRKMQLSLTSNKFSRKSSSYSSSST
VKTSY"

BASE COUNT 282 a 413 c 358 g 375 t

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:
1428	564.50	1428
Percent Similarity:	58.36%	Matches: 115
Best Local Similarity:	36.28%	Conservative: 70
Query Match:	26.43%	Mismatches: 121
		Indels: 12
DB:	10	Gaps: 5

US-09-208-629F-3 (1-407) x RNUG1373 (1-1428)

QY	95	VAlAsnAsnAlaThrIleGlyTyrLeuArgSerSerLeuSerThrGlnValIleProAla	114
DB	181	GTTGATGAATTCTCGCATCCGCCTCACGGGAAGCTGACCACCGCTCTTCCTCCCGGTC	240
QY	115	IleTyrlSeuLeuPheValValGlyValProSerAsnIleValThrLeuTrpHisLeu	134
DB	241	ATCTACATCATGCTCTTGTAATTGGTTGGCCAGTAATGGTATGGCCCTCTGGGCTTC	300
QY	135	SerLeuArgThr---LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla	153
DB	301	TTCCTCCGAACGAAGAAGAACCACTGCTGTGATTATACATGGCCAACCTGGCCTTGGCA	360
QY	154	AspLeuLeuPheCysValThrLeuProPheLyseIleAlaTyrlHisLeuAsnGlyAsnAsn	173
DB	361	GA CCTCTCTCTGTCATCTGGTTCCTCCCTGAAGATCTCTACCACTTCATGCCAAGCAC	420
QY	174	TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr	193
DB	421	TGGACCTATGGGATGGCTCTCAAGGTGCTCATTTGGCTTTTTTCTACGGCAATAATGTC	480
QY	194	CysAlaIleLeuIleLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro	213
DB	481	TGCTCCATCTTTTCATGACCTGCCCTCAGCGTCAGAGTACTGGGTGATCGTGAACCCC	540
QY	214	PheThrTyrglnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIleValTrp	233
DB	541	ATGGGACATCCAGG---AAGAGGGCAACATCGCTGTTGGCGTCTCCCTGGCCATCTGG	597
QY	234	ValMetValPheLeuTyrMetLeuProPheValIleLeuGlnGlutryHisLeuVal	253

598 CTCCTGATTTTCTGGTCCACCATCCTCTGTACGTATGAGGCACACCATCTACATTCCA 657

254 HisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerProSerSerPhe 273

658 GCCTTGAACATCATCACCCACCTGTGCAGACGTGCTGCCCGAG--GAGGTCTCTGGTGGGGAC 714

274 ArgPheTyrTyrPheValSerLeuAlaPhePheGlyPheLeuIlePheProPheValIlelle 293

715 ATGTTTCAGTTACTTCTCCTCCCTGCCCATTCGAGTCTTCTTCTGTTCCACGCCCTCCTTACT 774

294 IlePheCysTyrThrThrLeuIleHisIleCysLeuSerS-----Lys 307

775 GCGTCTGCTACGTGCTCATCATCAAAACGCTCCGCTCTTCGCGCATGAGCAGGACTCG 834

308 AspArgGluIleTrpLeuGlyTyrIleLeuValValLeuLeuIleValIlePheThrille 327

835 GAGAAGAAAAGCGGAGGGCTATCGCCTCATCATCAGCGTGTCTTCATGTACTTTCATC 894

328 CysPheAlaProThrAsnIleLeuValIleHisAlaAsnTyrTyrTyrHisAsn 347

895 TGCCTCGCTCCAGCAACGTGCTGCTGTCATTTCTCTCATCAAAAGCCAGAGG 954

348 ThrAspSerLeuTyrPheMetTyrIleLeuAlaLeuCysLeuGlySerLeuAsnSerCys 367

955 CAGAGCCACAGCTACGCGCTCTACCTCGTCCGCTCTGCTGTCCACCTCAACAGCTGC 1014

368 LeuAspProPheLeuTyrPheValMetSerIleValValAspGlnLeuAsnPro***Ser 387

1015 ATAGACCCCTTTGCTACTACTTTGTTTGGAAA-----GATTT-CAGGACCGAGCCNAG 1067

388 AlaMetAlaArgProLeu***ArgProArgArgAspIleTrpGluAspIle 404

1068 AAACGCGCTCTCTCGCGAAGCGTCCGACCGGTGAACGCATGCAGATATC 1118

RESULT 27 AC130634/c

LOCUS

DEFINITION

AC130634.1 GI:22212993

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Muzny,D.Marie, Metzker,M.Lee,, Abramzon,S., Adams,C., Alder,J., Allen,C., Allent,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Bryant,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bwalya,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregorgis,E., Gear,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M., Hollins,B., Howells,S., Huiyok,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowls,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhea,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarunpaga, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Plannkoc, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Ruiz, J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Wright, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 151166)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (13-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GXIX
 Center clone name: CH230-477A14
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 115058 bases at least Q40
 Consensus quality: 119558 bases at least Q30
 Consensus quality: 122439 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 37 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1149: contig of 1149 bp in length
 * 1150 1249: gap of unknown length
 * 1250 2920: contig of 1671 bp in length
 * 2921 3020: gap of unknown length
 * 3021 4429: contig of 1409 bp in length
 * 4430 4529: gap of unknown length
 * 4530 5676: contig of 1147 bp in length
 * 5677 5776: gap of unknown length
 * 5777 7247: contig of 1471 bp in length
 * 7248 7347: gap of unknown length
 * 7348 8743: contig of 1396 bp in length
 * 8744 8843: gap of unknown length
 * 8844 10237: contig of 1394 bp in length
 * 10238 10337: gap of unknown length
 * 10338 11514: contig of 1177 bp in length
 * 11515 11614: gap of unknown length

* 11615 12982: contig of 1368 bp in length
 * 12983 13082: gap of unknown length
 * 13083 14431: contig of 1349 bp in length
 * 14432 14531: gap of unknown length
 * 14532 17021: contig of 2490 bp in length
 * 17022 17121: gap of unknown length
 * 17122 19034: contig of 1913 bp in length
 * 19035 19134: gap of unknown length
 * 19135 20418: contig of 1283 bp in length
 * 20419 20517: gap of unknown length
 * 20518 22781: contig of 2264 bp in length
 * 22782 22881: gap of unknown length
 * 22882 25072: contig of 2191 bp in length
 * 25073 25172: gap of unknown length
 * 25173 26957: contig of 1785 bp in length
 * 26958 27057: gap of unknown length
 * 27058 29399: contig of 2342 bp in length
 * 29400 29499: gap of unknown length
 * 29500 33463: contig of 3964 bp in length
 * 33464 33563: gap of unknown length
 * 33564 35346: contig of 1783 bp in length
 * 35347 35446: gap of unknown length
 * 35447 38703: contig of 3257 bp in length
 * 38704 38803: gap of unknown length
 * 38804 42453: contig of 3650 bp in length
 * 42454 42553: gap of unknown length
 * 42554 45562: contig of 3009 bp in length
 * 45563 45662: gap of unknown length
 * 45663 49055: contig of 3403 bp in length
 * 49056 49165: gap of unknown length
 * 49166 51875: contig of 2710 bp in length
 * 51876 51976: gap of unknown length
 * 51977 54628: contig of 2653 bp in length
 * 54629 54729: gap of unknown length
 * 54730 58389: contig of 3661 bp in length
 * 58390 58489: gap of unknown length
 * 58490 62913: contig of 4424 bp in length
 * 62914 63013: gap of unknown length
 * 63014 68649: contig of 5636 bp in length
 * 68650 68749: gap of unknown length
 * 68750 76294: contig of 7545 bp in length
 * 76295 76394: gap of unknown length
 * 76395 81146: contig of 4752 bp in length
 * 81147 81246: gap of unknown length
 * 81247 86683: contig of 5437 bp in length
 * 86684 86783: gap of unknown length
 * 86784 91620: contig of 4837 bp in length
 * 91621 91720: gap of unknown length
 * 91721 99377: contig of 7657 bp in length
 * 99378 99477: gap of unknown length
 * 99478 108730: contig of 9253 bp in length
 * 108731 108830: gap of unknown length
 * 108831 120614: contig of 11784 bp in length
 * 120615 120714: gap of unknown length
 * 120715 130786: contig of 10072 bp in length
 * 130787 130886: gap of unknown length
 * 130887 151166: contig of 20280 bp in length.
 FEATURES
 Location/Qualifiers
 1. .151166
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /clone="CH230-477A14"
 BASE COUNT 37400 a 32808 c 33338 g 36965 t 10655 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.24e-34 Length: 151166
 Score: 564.50 Matches: 115
 Percent Similarity: 58.36% Conservative: 70
 Best Local Similarity: 36.28% Mismatches: 121
 Query Match: 26.43% Indels: 12
 DB: 2 Gaps: 5

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Db 43849 ATTGCCTATCACATACATACGCAACACTGGATTATATGGGAAGCTCTTTTGTAAATGCTT 43908
 Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuThrCysMetGlyIle 204
 Db 43909 ATTGGCTTTTCTATGCAACATGACTGTTCCATTCTCTCATGACCTGCTCAGTGTG 43968
 Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
 Db 43969 CAGAGGTATTGGGTCACTGTCGAACCCCATGGGCACCTCCAGG--AAGAAGGCAAAACATT 44025
 Qy 225 SerLeuLeuMetCysGlyIleValTrrpValMetValPheLeuTyrMetLeuProPheVal 244
 Db 44026 GCATTGGCATCTCCCTGGCAATATGCTGCTGATCTGCTGTCACCATCTCTTGTAT 44085
 Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrCysHisAspValVal 264
 Db 44086 GTCGTGAAGCAGACCATCTTCAATCTCTGCTGCAACATCAGACCTGTCATGATGTTTG 44145
 Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
 Db 44146 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCT 44190
 Qy 281 LeuAlaPhePheGlyPheLeuIlePheProPheValIleIlePheCysTyrThrThrLeu 300
 Db 44191 CTGCGCAATGGGTCTTTCTGTTCCCGACCTTCTCAGCCCTGCTGCTATGCTGATG 44250
 Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTrrpLeuGlyTyr 314
 Db 44251 ATCAGAATGCTGGATCTTCTGCCATGATGAACTCAGAGAAGAAAGAGAGGGCC 44310
 Qy 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 44311 ATCAAACTATTGTCACTGCTGGCCATGTACCTGATCTGCTCACTCTAGTAACCTT 44370
 Qy 335 IleLeuValIleHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
 Db 44371 CTGCTTGTGTGCAATATTTCTGATTAAAGCAGCCAGGCGCAGCATGCTATGCCCTG 44430
 Qy 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
 Db 44431 TACATTGAGCCCTCTGCTCTCTACCTTAAAGCTGATCGACCCCTTGTCTATTAC 44490
 Qy 375 ValMetSer 377
 Db 44491 TTTGTTTCA 44499

RESULT 29

AR012638 LOCUS 1255 bp DNA linear PAT 04-DEC-1998
 DEFINITION Sequence 3 from patent US 5763575.
 ACCESSION AR012638
 VERSION AR012638.1 GI:3970628
 KEYWORDS
 SOURCE Unknown.

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 1255)

Sundelin, J. and Scarborough, R.M.

Agonist and antagonist peptides of the C140 receptor

JOURNAL Patent: US 5763575-A 3 09-JUN-1998;

Location/Qualifiers

source 1..1255

/organism="unknown"

BASE COUNT 294 a 320 c 260 g 381 t

ORIGIN

Alignment Scores:

Pred. No.: 7.67e-37 Length: 1255
 Score: 563.00 Matches: 115
 Percent Similarity: 54.52% Conservativeness: 72
 Best Local Similarity: 33.53% Mismatches: 128
 Query Match: 26.36% Indels: 28
 DB: 6 Gaps: 7

US-09-208-629f-3 (1-407) x AR012638 (1-1255)
 Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
 Db 116 ACCCTTGCTCTCTTCTTGTACAGGAACCAATAGATCTCTAAAGGAAGAACCTTATT 175
 Qy 69 SerAspIleGluGlyTrrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
 Db 176 GGTAAAGTTGATGGCAGCATCCCAACGTCACCTGGAAAAAGGAGTTACAGTTCAA----- 226
 Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
 Db 227 -----ACAGTCTTTTCTGTGGATGAGTTTCTGTCATCTGTCTCTCACTGGA 271
 Qy 106 SerLeuSerThrGlnValIleProAlaIleTrrpIleLeuLeuPheValValGlyValPro 125
 Db 272 AAATGACCACTGCTCTCTCCATTGCTACAAATGTTGTTGTGGTGGTGGTGGCC 331
 Qy 126 SerAsnIleValThrLeuTrrpLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
 Db 332 AGTAACGGCATGCCCTGTGGGTCTTCTTTCCGAACCTAAGAAGAACGACCCCTGCTGTG 391
 Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
 Db 392 ATTACATGCCCAATCTGGCTTGGCTGACCTCTCTCTGTCATCTGCTGTTCCCTTGAAG 451
 Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrrpValPheGlyGluValMetCysArgIleThr 184
 Db 452 ATTGCTATCATATACATACGCAACAACTGGATTATGGGGAAGCTCTTTGTAATGCTT 511
 Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
 Db 512 ATTGGCTTTTCTATGGCAACATGACTGTTCCATTCTCTCATGACCTGCCTCAGTGTG 571
 Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
 Db 572 CAGAGGTATTGGGTCACTGTCGAACCCCATGGGCACCTCCAGG--AAGAAGGCAAAACATT 628
 Qy 225 SerLeuLeuMetCysGlyIleValTrrpValMetValPheLeuTyrMetLeuProPheVal 244
 Db 629 GCATTGGCATCTCCCTGGCAATATGCTGCTGATTTCTGCTGTCACCATCTCTTGTAT 688
 Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
 Db 689 GTCGTGAAGCAGACCATCTTCAATCTCTGCTGCAACATCAGACCTGTCATGATGTTTG 748
 Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
 Db 749 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCTCT 793
 Qy 281 LeuAlaPhePheGlyPheLeuIlePheProPheValIleIlePheCysTyrThrThrLeu 300
 Db 794 CTGCGCAATGGGTCTTCTTGTTCACGCTCTCTCAGCCCTCTCCTCATGCTGCTGATG 853
 Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTrrpLeuGlyTyr 314
 Db 854 ATCAGAATGCTGGATCTTCTGCCATGATGAAACTCAGAGAAGAAAGAGAGAGGGCC 913
 Qy 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 914 ATCAAACTCATTTGCTCACTGCTGCGCATGTACTGCTTCACTCTGCTTAGTAACCTT 973
 Qy 335 IleLeuValIleHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
 Db 974 CTGCTTGTGTGCAATATTTTCTGATTAAAGCAGCCAGGCGCAGACCATGCTATGCCCTG 1033
 Qy 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
 Db 1034 TACATTGAGCCCTCTGCTCTCTACCTTAAAGCTGATCGACCCCTTGTCTATTAC 1093
 Qy 375 ValMetSer 377
 Db 375 ValMetSer 377

Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Db 272 AAACCTGACCACTGCTTCCTTCCAAATGCTACACAATGTTGTTGGTGGGTTTCCCA 331
Qy 126 SerAsnIleValThrLeuTyrLysLeuSerLeuArgThr--LysSerIleSerLeuVal 144
Db 332 AGTAACGGCATGGCCCTTGCGTCTTCTTCCGAACCTAAGAAAGAGCACCCTGCTGTG 391
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 392 ATTTACATGGCAATCTGGCCTTGACCTCTCTCTGTCATCTGTTCCCTTGAAG 451
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 452 ATTGCCATCATGCTGCTGGCCATGACCTGATCTGCTTCACTCTGCTAGTAACCT 973
Qy 335 IleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
Db 974 CTGCTTGTGGTGCATTTATTTCTGATTAAGAGCCAGCCAGCCATGCTATGCCCTG 1033
Qy 375 ValMetSer 377
Db 1094 TTTGTTTCA 1102
RESULT 32
187849
LOCUS 1255 bp DNA linear PAT 10-AUG-1998
DEFINITION Sequence 3 from patent US 5716789..
ACCESSION 187849
VERSION 187849.1 GI:3407789
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1255)
AUTHORS Sundelin,J. and Scarborough,R.M.
TITLE Method to determine ligands, agonist and antagonist of C140

receptor
Patent: US 5716789-A 3 10-FEB-1998;
FEATURES Location/Qualifiers
source 1. .1255
BASE COUNT 294 a 320 c 260 g 381 t
ORIGIN
Alignment Scores:
Pred. No.: 7,67e-37 Length: 1255
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 6 Gaps: 7
US-09-208-629F-3 (1-407) x 187849 (1-1255)
Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 116 ACCCTTGTCTTCTTCTTCTTGTACAGGAACCAATAGATCTCTTAAGAGGAAGCCTTATT 175
Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrThrIleLysAlaGluCys 85
Db 176 GGTAAAGTTGATGGCACAATCCACGCTACCTGAGAAAGAGGATTACAGTTGAA----- 226
Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 227 -----ACAGTCTTTTGTGGATGAGTTTCTGCATCTGCTCCTCCTGGA 271
Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro 125
Db 272 AAACCTGACCACTGCTTCCTTCCAAATGCTACACAATGTTGTTGGTGGGTTTCCCA 331
Qy 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr--LysSerIleSerLeuVal 144
Db 332 AGTAACGGCATGGCCCTTGCGTCTTCTTCCGAACCTAAGAAAGAGCACCCTGCTGTG 391
Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 392 ATTTACATGGCAATCTGGCCTTGACCTCTCTCTGTCATCTGTTCCCTTGAAG 451
Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 452 ATTGCCATCATGCTGCTGGCCATGACCTGATCTGCTTCACTCTGCTAGTAACCT 511
Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyLe 204
Db 512 ATTGGCTTTTCTATGGCAACATGCTGTTCCATTTCTTCTATGACCTGCCCTCAGGTG 571
Qy 205 AspArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 572 CAGAGGTATGGGTCTATCTGAACCCCATGGGGCCTCCAGG---AAGAGGCAACAT 628
Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 629 GCCATTGGCATCTCCCTGGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 688
Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 689 GTCGTGAAGCAGACCATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
Qy 265 AspAlaCysGluSerProSerPhe-----ArgPheTyrTyrPheValSer 280
Db 749 -----CCTGAGCAGCTCTTGTGGGAGACATGTTTCAATTACTTCTCTCT 793
Qy 281 LeuAlaPhePheGlyPheLeuProPheValIleIleIlePheCysTyrThrThrLeu 300
Db 794 CTGGCCATTTGGGGTCTTCTGTTCCAGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 853
Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 854 ATCAGATGCGCGCATCTTCTGCCATGATGAATACTCAGAGAGAAAGAGAGAGGCC 913
Qy 315 IleLysAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 914 ATCAAACTCATGTGCTGCTGGCCATGACCTGATCTGCTTCACTCTGCTAGTAACCT 973
Qy 335 IleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
Db 974 CTGCTTGTGGTGCATTTATTTCTGATTAAGAGCCAGCCAGCCATGCTATGCCCTG 1033
Qy 355 TyrLeuIleAlaLeuLysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
Db 1034 TACATTGTAGCCCTTCTGCTCTACCCCTTAACAGCTGATCGACCCCTTGTCTATTAC 1093
Qy 375 ValMetSer 377
Db 1094 TTTGTTTCA 1102
RESULT 32
187849
LOCUS 1255 bp DNA linear PAT 10-AUG-1998
DEFINITION Sequence 3 from patent US 5716789..
ACCESSION 187849
VERSION 187849.1 GI:3407789
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1255)
AUTHORS Sundelin,J. and Scarborough,R.M.
TITLE Method to determine ligands, agonist and antagonist of C140

Qy		315	IleLeuAlaValLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle	334
			:::::	
Db		914	ATCAAACTCATGTGTCCTGGCCATGTACCTGCATTCTCCTAGTAACCTT	973
			:::::	
Qy		335	IleLeuValIleHisAlaAsnTyrrTyrrHisasnThrAspSerLeuTyrrPheMet	354
			:::::	
Db		974	CTGCTTGTTGGTCATTATTTCTGATTGAAGCCAGGCCCATGCTATGTCCTG	1033
			:::::	
Qy		355	TyrLeuileAlaLeuCysLysLeuAsnSercYsisLeuAspPropheLeuTyrrPhe	374
			:::::	
Db		1034	TACATTGTAGCCCTCTGCCTCTCTACCCCTAACAGCTGCATCGACCCTTGTCTATTAC	1093
			:::::	
Qy		375	ValMetSer	377
			:::::	
Db		1094	TTTGTTCACA	1102

RESULT	33
HSPAR2B	
LOCUS	HSPAR2B
DEFINITION	H.sapiens partial gene for proteinase-activated receptor 2 (1789 BP).
	1789 bp DNA linear PRI 21-JUL-1997

US-09-208-629F-3 (1-407) x HSPAR2B (1-1289)

Qy	49	ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu	68
Db	72	ACCCTTGCTCTCTTCTTGTCACGAAACCAATAGATCCTCTAAAGGAAGAGCCTTATT	131
Qy	69	SerAspIleGluGlyTyr-----ThrGlyAlaThrThrThrIleLysAlaGluCys	85
Db	132	GGTAAGGTTGTATGGACATCCACATCCCACTGCTGAAAGGAGGTTTACAGTTGAA-----	182

Qy	86	ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer	105
Db	183	-----ACAGTCCTTTCTGGTGGATGAGTCTTCTGCATCTGTCCTCACTGGGA	227
Qy	106	SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValPro	125
Db	228	AAACTGACCACCTGCTTCCTTCCAATGTCTACAAATTGTGTGTGGTGGGTGTGCCA	287
Qy	126	SexAsnIleValThrLeuTyrLysLeuSerLeuArgThr---LysSerIleSerLeuVal	144
Db	288	AGTAACGGCATGCCCTGTGGTCTTCITTTCCGAAGTAAGAAGAACGCCCTGCTGTG	347
Qy	145	IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys	164
Db	348	ATTTCATGCGCAATCTGGCTTGGGTGACCTCTCTCTGTCATCTGGTTCCCTCTGAAG	407
Qy	165	IleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValMetCysArgIleThr	184
Db	408	ATTGCCCTATCACATACATGACCAACTGGAATTTATGGGAAGCTCTTTGTAATGTGCTT	467
Qy	185	ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThrCysMetGlyIle	204
Db	468	ATTGGCTTTTCTATGGCAACATGACTGTGTCCATCTCTTCATGACCTGCCCTCAGGTG	527
Qy	205	AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe	224
Db	528	CAGAGGTATTGGGTCATCGTGAACCCCATCGGGGCATCCACGG---AAGAAGGCAAACTT	584
Qy	225	SerLeuLeuMetCysGlyIleValTyrPvalMetValPheLeuTyrMetLeuProPheVal	244
Db	585	GCCATTGGCATCTCCCTGGCAATATGGCTGCTGATCTGCTGGTCACCATCCCTTTGAT	644
Qy	245	IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal	264
Db	645	GTCGTGAAGCAGACCATCTTCATCTCCTGCGCCTGAACATCACGACCTGTCATGATGTTT	704
Qy	265	AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer	280
Db	705	-----CTGAGCAGCTCTTGGTGGGAGACATGTTCATTAATCTCTCTCTCTCT	749
Qy	281	LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu	300
Db	750	CTGGCCATTGGGGTCTTTCTGTTCCAGCCTTCTCTCACAGCCTCTGCTATGCTGCTGATG	809
Qy	301	IleHisLysLeuLysSer-----LysAspArgIleThrLeuGlyTyr	314
Db	810	ATCAGAATGTGCGCATCTTCTGCCATGGATGAAACTCAGAGAAGAAAGAAAGAGAGGCC	869
Qy	315	IleLysAlaValLeuLeuLeuLeuValIlePheThrIleCysPheAlaProThrAsnIle	334
Db	870	ATCAAACTCATGTGTCATGCTCGGCCCATGACTGATCTGCTTCATCTCCTAGTAACCTT	929
Qy	335	IleLeuValIleHisHisAlaAsnTyrTyrHisAsnThrAspSerLeuTyrPheMet	354
Db	930	CTGCTTGTGGTCATTATTCTTGATTAGACGCCAGGGCCAGAGCCATGCTATGCCCTG	989
Qy	355	TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe	374
Db	990	TACATTGTAGCCCTCTGCGCTCTCTACCCCTTAACAGCTGCATCGACCCCTTTGTCTATTAC	1049
Qy	375	ValMetSer	377
Db	1050	TTTGTTTCA	1058
RESULT 34			
HSU34038			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
Homo sapiens			
1451 bp mRNA linear PRI 21-FEB-1997			
Human proteinase-activated receptor-2 mRNA, complete cds.			

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1451)
 Bohm S.K., Kong W., Bromme D., Smeekens S.P., Anderson D.C.,
 Connolly A., Kahn M., Neiken N.A., Coughlin S.R., Payan D.G. and
 Bunnett N.W.
 TITLE Molecular cloning, expression and potential functions of the human
 proteinase-activated receptor-2
 JOURNAL Biochem. J. 314 (Pt 3), 1009-1016 (1996)
 MEDLINE 96177879
 PUBMED 8615752
 REFERENCE 2 (bases 1 to 1451)
 Bohm S.
 TITLE Direct Submission
 JOURNAL Submitted (15-AUG-1995) Stephan Bohm, Department of Surgery, School
 of Medicine, University of California at San Francisco, 521
 Parnassus Avenue, San Francisco, CA 94143-0660, USA
 FEATURES
 Location/Qualifiers
 1..1451
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="kidney"
 1..147
 148..1341
 /note="G-protein-coupled receptor"
 /codon_start=1
 /product="proteinase-activated receptor-2"
 /protein_id="AA84781.1"
 /db_xref="GI:1041729"
 /translation="MRSPAAWLLGAILAALSLSGTTQGNRRSKGRSLGKVDG
 TSHVGTGVTETVFSVDFPSALVTKLTPIVTVFVVGSLPSNGMALWFLF
 RTKKHPAVLYAMNALLSLVWFKLKIAYHANNWYIGALCNLVLPYFGYNNY
 CSTLFWCLSVORWTVNPMWHSRKKANIAIGISLAIWLLILVTPLVVKOTIFI
 PALNITTHCVLPQLVGMFNYFLSLAIGVLPFPAFLTASAYVLMIRLSANDE
 NSKKRRALKLVTLVLAFLICFTPSNLVLLVHYFLDKSQGSHVLAFLVALCLST
 LNSCIDPFVYFVSHDFRDAKNALLCRSVRTKQMVSLTSKKHRSKSSSSSTT
 VKTSY"
 1342..1451
 310 a 389 c 346 g 406 t
 BASE COUNT
 ORIGIN
 Alignment Scores:
 Pred. No.: 9.02e-37 Length: 1451
 Score: 563.00 Matches: 119
 Percent Similarity: 52.32% Conservative: 73
 Best Local Similarity: 32.43% Mismatches: 137
 Query Match: 26.36% Indels: 8
 DB: 9 Gaps: 8
 US-09-208-629f-3 (1-407) x HSU34038 (1-1451)
 Qy 22 LeuValAlaAlaGlyLeuLeuPheLeuProValThrValCysGlnSerGlyIleAsnVal 41
 Db 172 CTGCTGGGGGCCCATCTCTGCTAGCAGCCTCTCTCTGCGATGGCCACCATCCAGGA 231
 Qy 42 SerAsnSerAlaLeuLysProThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsn 61
 Db 232 ACCAATAGATCCTCTAAGGAGAGCGCTTATGGTAGGTGATGC----- 279
 Qy 62 ThrPheGluGluPheProLeuSerAspIleGluGlyTrpThrGlyAlaThrThrIle 81
 Db 280 -----ACATCCACGTC-----ACTGGAAGAGGATTACAGTT 312
 Qy 82 LysAlaGluCysProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGly 101
 Db 313 GAA-----ACAGTCTTTCTGCGATGAGTTTCTGCATCT 348
 Qy 102 TyrLeuArgSerSerLeuThrGlnValIleProAlaIleTyrIleLeuLeuPheVal 121
 Db 349 GTCTCTACTGGAACCTGACCACCGTCTCTCCATTTGCTACAAATTTGTTTGTG 408
 Qy 122 ValGlyValProSerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSer 140

Db 409 GTGGGTTTCCCAAGTAACGGCATGGCCCTGTGGGTCTTTCTTTCGAACTAAGAAGAAG 468
 Qy 141 IleSerLeuValIlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThr 160
 Db 469 CACCCTGCTGTGATTACATGGCCATCTGGCTTGCTGACCTCTCTCTGTCATCTGG 528
 Qy 161 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMet 180
 Db 529 TTCCCTTGAAGATTGCCTATCACATACATGCACTGGGATTTATGGGAAGCTCTT 588
 Qy 181 CysArgIleThrThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuThr 200
 Db 589 TGTAAATGCTGCTATTGGCTTTTCTATGGCAACATGCTACTGTTCCTTCTTCATGACC 648
 Qy 201 CysMetGlyIleAsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuPro 220
 Db 649 TGCTCAGTGTGCAGAGGATTGGGTGCATCGTGAACCCCATGGGGCAGCTCCAGG---AAG 705
 Qy 221 LysArgSerPheSerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMet 240
 Db 706 AAGGCAACATTCGCCATTCCTCCCTGGCAATATGCTGTGATTTCTGCTGGTCACC 765
 Qy 241 LeuProPheValIleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCys 260
 Db 766 ATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATTCCTGCTGCAACATCAGCAGCTGT 825
 Qy 261 HisAspValValAspAlaCysGluSerProSerSerPhe-----ArgPheTyr 276
 Db 826 CATGATGTTTGT-----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 870
 Qy 277 TyrPheValSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCys 296
 Db 871 TACTTCTCTCTGCGCATTCGGGCTTTCTGTTCCAGCCTTCTCTCAGAGCTCTGCC 930
 Qy 297 TyrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
 Db 931 TATGTCGTGATGATGAGAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAAGAA 990
 Qy 311 TrpLeuGlyTyrIleLysAlaValLeuLeuIlePheThrIleCysPheAla 330
 Db 991 AGGAAGAGGCCATCAACTCATTTGCTGCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCT 1050
 Qy 331 ProThrAsnIleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
 Db 1051 CCTAGTAACCTTCTGCTTGTGGTGCAATTTTCTGATTAAAGAGCCAGGSCCAGAGCAT 1110
 Qy 351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
 Db 1111 GTCTATGCCCTGTACATTGTAGCCCTCTGCTCTCTACCCCTTAACAGCTGCATCGACCCC 1170
 Qy 371 PheLeuTyrPheValMetSer 377
 Db 1171 TTGTCTATTACTTTGTTTCA 1191
 RESULT 35
 AF400075 18351 bp DNA linear PRI 18-JAN-2002
 LOCUS Homo sapiens coagulation factor II (thrombin) receptor-like 1
 DEFINITION (F2RL1) gene, complete cds.
 ACCESSION AF400075
 VERSION AF400075.1 GI:15021772
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 18351)
 AUTHORS Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L.,
 Yi Q. and Nickerson D.A.
 TITLE Direct Submission
 JOURNAL Submitted (17-JUL-2001) Molecular Biotechnology, University of
 Washington, 1705 NE Pacific, Seattle, WA 98195, USA

```

COMMENT      To cite this work please use:  SeattleSNPs, NHLBI Program for
Genomic Applications, UW-PHCRC, Seattle, WA (URL:
http://pga.mbt.washington.edu)
This sequence consists of 2 contigs.  The gap between the contigs
is represented as a run of N.
1 10774: contig of 10774 bp in length
10775 10874: gap of unknown length
10875 18351: contig of 7477 bp in length.
FEATURES
source      1..18351
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
repeat_region 343..492
            /rpt_family="MIR"
            /rpt_type="dispersed"
variation    467
            /frequency="0.09"
            /replace="a"
variation    531
            /frequency="0.01"
            /replace="g"
variation    678
            /frequency="0.01"
            /replace="c"
repeat_region 728..780
            /rpt_family="L2"
            /rpt_type="dispersed"
variation    829
            /frequency="0.03"
            /replace="a"
repeat_region 913..1028
            /rpt_family="FLAM_C"
            /rpt_type="dispersed"
variation    925
            /frequency="0.01"
            /replace="t"
repeat_region 1059..1214
            /rpt_family="L3"
            /rpt_type="dispersed"
repeat_region 1292..1578
            /rpt_family="AluSg"
            /rpt_type="dispersed"
variation    1538
            /frequency="0.01"
            /replace="g"
variation    1622
            /frequency="0.02"
            /replace="t"
variation    1623
            /frequency="0.01"
            /replace="g"
repeat_region 1795..1926
            /rpt_family="MIR"
            /rpt_type="dispersed"
variation    2122
            /frequency="0.01"
            /replace="g"
variation    2254
            /frequency="0.01"
            /replace="c"
variation    2457
            /frequency="0.10"
            /replace="g"
gene         2462..17176
            /gene="F2RL1"
mRNA         join(2462..2687,15958..17176)
            /gene="F2RL1"
            /product="coagulation factor II (thrombin) receptor-like
            1"
variation    2561
            /gene="F2RL1"
            /frequency="0.28"
            /replace="c"
CDS
join(2606..2687,15958..17069)
/gene="F2RL1"
/codon_start=1
/product="coagulation factor II (thrombin) receptor-like
1"
/protein_id="AAK77914.1"
/db_xref="GI:15021773"
/translation="MRSPSAWLLGAAIILAAASLCSGTTIOGTNRSSKGRSLIGKVDG
TSHVTGKGVTVETVSDEFSASVLTGKLTTVFLPVTIVTVFVGLPSNGMALWVFLF
RYKKHPAVIYMANALADLLSVINFPKIAIHIGNNWIYGEALCNVLIGFYGNMY
CSILFMTCLSVORYWVINPMSGHRSRKANIAIGISLAIMLLILLVLTPIYVVKQTIFI
PALNITTDHDLVPEQLVGDMEFYFLSAIGVFLPFAFLTASAYVLMIRSSAMDE
NSEKKRKAIKLITVLAWYLICTPSPNLLLVVHYFLIKSQSQSHVYALYIVALCLST
LNSCIDPFVYFVSHDFRDHAKNALLCRSVRTVKQMVSLTSKHSRKSSYSSTT
VKTSY"
variation    2667
            /gene="F2RL1"
            /frequency="0.02"
            /replace="t"
variation    2770
            /gene="F2RL1"
            /frequency="0.09"
            /replace="a"
variation    3069
            /gene="F2RL1"
            /frequency="0.03"
            /replace="a"
variation    3119
            /gene="F2RL1"
            /frequency="0.32"
            /replace="t"
variation    3151
            /gene="F2RL1"
            /frequency="0.01"
            /replace="g"
variation    3263
            /gene="F2RL1"
            /frequency="0.33"
            /replace="g"
variation    3306
            /gene="F2RL1"
            /frequency="0.02"
            /replace="t"
variation    3330
            /gene="F2RL1"
            /frequency="0.01"
            /replace="c"
variation    3331
            /gene="F2RL1"
            /frequency="0.07"
            /replace="t"
variation    3353
            /gene="F2RL1"
            /frequency="0.01"
            /replace="a"
variation    3417
            /gene="F2RL1"
            /frequency="0.01"
            /replace="c"
variation    4160
            /gene="F2RL1"
            /frequency="0.01"
            /replace="c"
variation    4456
            /gene="F2RL1"
            /frequency="0.01"
            /replace="t"
variation    4500
            /gene="F2RL1"
            /frequency="0.02"
            /replace="t"
variation    4663
            /gene="F2RL1"

```

```

/replacement="0.46"
/replacement="t"
4710.4817
/rpt_family="L2"
/rpt_type="dispersed"
4715
/gene="F2RL1"
/frequency="0.15"
/replacement="g"
4972
/gene="F2RL1"
/frequency="0.02"
/replacement="c"
4993.5299
/rpt_family="AluSx"
/rpt_type="dispersed"
5065
/gene="F2RL1"
/frequency="0.01"
/replacement="g"
5159
/gene="F2RL1"
/frequency="0.01"
/replacement="a"
5295
/gene="F2RL1"
/frequency="0.01"
/replacement="t"
5296
/gene="F2RL1"
/frequency="0.01"
/replacement="a"
5366
/gene="F2RL1"
/frequency="0.17"
/replacement="t"
5643.5940
/rpt_family="AluSx"

```

Alignment Scores:

```

Pred. No.: 1.55e-35 Length: 18351
Score: 563.00 Matches: 115
Percent Similarity: 54.52% Conservative: 72
Best Local Similarity: 33.53% Mismatches: 128
Query Match: 26.36% Indels: 28
DB: 9 Gaps: 7

```

US-09-208-629F-3 (1-407) x AF400075 (1-18351)

```

Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
Db 15933 ACCCTTGCTTCCTTCTTGACAGGAACCAATAGATCCTCTAAAGGAAGAGCCTTATT 15992

Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrIleLysAlaGluCys 85
Db 15993 GGTAAAGTTGATGCACATCCCGTCAGTGAAGAGGATTACAGTTGAA----- 16043

Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAlaThrIleGlyTyrLeuArgSer 105
Db 16044 -----ACAGCTTTTCTGTGATGAGTTTCTGCATCTGCTCCTCCTGGA 16088

Qy 106 SerLeuSerThrGlnValIleProAlaIlePheLeuLeuPheValValGlyValPro 125
Db 16089 AAACCTGACCACTGCTTCCTCCCAATGCTACACAATGTTGTTGTTGGTGGTTGCCA 16148

Qy 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
Db 16149 AGTAACGCGCATGGCCCTGTGGGTCTTTCTTCCGAACTAAGAAAGACACCCCTGCTGTG 16208

Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuLeuPheCysValThrLeuProPheLys 164
Db 16209 ATTTACATGGCAATCTGGCCTTGGCTGACCTCTCTCTGTCATCTGGTTCCTTGAAG 16268

```

```

Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
Db 16269 ATTGCCTATCATATACATGCAACAACTGATTTATGGGAGAGCTCTTTGTAATGCTT 16328

Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuLeuLeuLeuThrCysMetGlyIle 204
Db 16329 ATTGGCTTTTCTATGGCAACATGATGTTCTTCATTTCTTTCATGACCTGCCTCAGTGTG 16388

Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
Db 16389 CAGAGGTATTGGTGCATCGTGAACCCCTGAGGCACTCCAGG--AAGAAGGCAACATT 16445

Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
Db 16446 GCCATTGGCATCTCCCTCGCAATATGGCTGCTGATTTCTGCTGCTCACCCTCCCTTGTAT 16505

Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
Db 16506 GTCGTAGAGCAGACCATCTTCATCTGCTGCTGCAACATCAGACCTGTGATGATTTTG 16565

Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
Db 16566 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCTCT 16610

Qy 281 LeuAlaPhePheGlyPheLeuLeuPhePheValIleIlePheCysTyrThrThrLeu 300
Db 16611 CTGGCCATTGGGCTCTTCTGTTCCAGGCTTCTCCTCAGACCTCTGCTGCTGCTGATG 16670

Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
Db 16671 ATCAGAACTGCTGGATCTTCTGCTGATGAAACTCAGAGAGAAAGAGAGAGGGCC 16730

Qy 315 IleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAlaProThrAsnIle 334
Db 16731 ATCAAACTCATGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 16790

Qy 335 IleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
Db 16791 CTGCTTGTGTCATTTATTTCTGATTAAAGAGCAGGCGCAGAGCCATGCTATGCCCTG 16850

Qy 355 ThrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuPhe 374
Db 16851 TACATTGTAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 16910

Qy 375 ValMetSer 377
Db 16911 TTTGTTTCA 16919

RESULT 36
AC068682/c 184536 bp DNA linear HTG 23-SEP-2000
LOCUS
DEFINITION Homo sapiens clone RP11-206N2, WORKING DRAFT SEQUENCE, 37 unordered
pieces
AC068682
VERSION AC068682.3 GI:10280868
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 184536)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens, clone RP11-206N2
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 184536)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choe, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, W., Doyle, W., Ferreira, P., FitzHugh, W., Gage, D.,
Gallagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,

```

Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lanazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Sep 23, 2000 this sequence version replaced gi:8247861.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center Project name: L5802

Center clone name: 206_N2

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly: Phrap; version 0.960731

Consensus quality: 167659 bases at least Q40

Consensus quality: 175736 bases at least Q30

Consensus quality: 178914 bases at least Q20

Insert size: 154000; agarose-fp

Quality coverage: 5.0 in Q20 bases; agarose-fp

Quality coverage: 4.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 37 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 1154: contig of 1154 bp in length
 * 1155 1254: gap of 100 bp
 * 1255 2766: contig of 1512 bp in length
 * 2767 2866: gap of 100 bp
 * 2867 4073: contig of 1207 bp in length
 * 4074 4173: gap of 100 bp
 * 4174 5191: contig of 1018 bp in length
 * 5192 5291: gap of 100 bp
 * 5292 7387: contig of 2096 bp in length
 * 7388 7487: gap of 100 bp
 * 7488 9087: contig of 1600 bp in length
 * 9088 9187: gap of 100 bp
 * 9188 10361: contig of 1174 bp in length
 * 10362 10461: gap of 100 bp
 * 10462 12065: contig of 1604 bp in length
 * 12066 12165: gap of 100 bp
 * 12166 13498: contig of 1333 bp in length
 * 13499 13598: gap of 100 bp
 * 13599 15085: contig of 1487 bp in length
 * 15086 15185: gap of 100 bp
 * 15186 17249: contig of 2064 bp in length
 * 17250 17349: gap of 100 bp
 * 17350 19243: contig of 1894 bp in length
 * 19244 19343: gap of 100 bp
 * 19344 21127: contig of 1784 bp in length

* 21128 21227: gap of 100 bp
 * 21228 23345: contig of 2118 bp in length
 * 23346 23445: gap of 100 bp
 * 23446 25965: contig of 2520 bp in length
 * 25966 26065: gap of 100 bp
 * 26066 28195: contig of 2130 bp in length
 * 28196 28295: gap of 100 bp
 * 28296 32216: contig of 3921 bp in length
 * 32217 32316: gap of 100 bp
 * 32317 36954: contig of 4638 bp in length
 * 36955 37054: gap of 100 bp
 * 37055 40348: contig of 3294 bp in length
 * 40349 40448: gap of 100 bp
 * 40449 43556: contig of 3108 bp in length
 * 43557 43656: gap of 100 bp
 * 43557 47926: contig of 4270 bp in length
 * 47927 48026: gap of 100 bp
 * 48027 54079: contig of 6053 bp in length
 * 54080 54179: gap of 100 bp
 * 54180 58962: contig of 4783 bp in length
 * 58963 59062: gap of 100 bp
 * 59063 63739: contig of 4677 bp in length
 * 63740 63839: gap of 100 bp
 * 63840 71551: contig of 7712 bp in length
 * 71552 71651: gap of 100 bp
 * 71652 78218: contig of 6567 bp in length
 * 78219 78318: gap of 100 bp
 * 78319 84982: contig of 6664 bp in length
 * 84983 85082: gap of 100 bp
 * 85083 91930: contig of 6848 bp in length
 * 91931 92030: gap of 100 bp
 * 92031 98922: contig of 6892 bp in length
 * 98923 99022: gap of 100 bp
 * 99023 106496: contig of 7474 bp in length
 * 106497 106596: gap of 100 bp
 * 106597 114950: contig of 8354 bp in length
 * 114951 115050: gap of 100 bp
 * 115051 124353: contig of 9303 bp in length
 * 124354 124453: gap of 100 bp
 * 124454 132639: contig of 8186 bp in length
 * 132640 132739: gap of 100 bp
 * 132740 141572: contig of 8833 bp in length
 * 141573 141672: gap of 100 bp
 * 141673 151302: contig of 9630 bp in length
 * 151303 151402: gap of 100 bp
 * 151403 161767: contig of 10365 bp in length
 * 161768 161867: gap of 100 bp
 * 161868 184536: contig of 22669 bp in length.

FEATURES

Location/Qualifiers
 1. .184536
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="RP11-206N2"
 /clone_lib="RP11-206N2"
 1. .1154
 /note="assembly_fragment"
 clone_end:Sp6
 vector_side:left
 misc_feature
 1255..2766
 /note="assembly_fragment"
 2867..4073
 /note="assembly_fragment"
 4174..5191
 /note="assembly_fragment"
 5292..7387
 /note="assembly_fragment"
 7488..9087
 /note="assembly_fragment"
 9188..10161
 /note="assembly_fragment"
 10462..12065
 /note="assembly_fragment"
 12166..13498
 /note="assembly_fragment"

Consensus quality: 199004 bases at least Q30
 Consensus quality: 200415 bases at least Q20
 Estimated insert size: 175000; agarose-fp estimation
 Quality coverage: 9.78 in Q20 bases; agarose-fp estimation.
 Quality coverage: 8.4 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1058: contig of 1058 bp in length
 * 1059 1158: gap of unknown length
 * 1159 2342: contig of 1184 bp in length
 * 2343 2442: gap of unknown length
 * 2443 4907: contig of 2465 bp in length
 * 4908 5007: gap of unknown length
 * 5008 19888: contig of 14878 bp in length
 * 19889 19888: gap of unknown length
 * 19888 35850: contig of 15865 bp in length
 * 35851 35950: gap of unknown length
 * 35951 53802: contig of 17852 bp in length
 * 53803 53902: gap of unknown length
 * 53903 107788: contig of 53886 bp in length
 * 107789 107888: gap of unknown length
 * 107889 204485: contig of 96597 bp in length.

FEATURES

Location/Qualifiers
 1. 204485
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-206N2"
 /clone lib="RPC1 human BAC library 11"
 BASE COUNT 56752 a 44581 c 44954 g 57495 t 700 others
 ORIGIN

Alignment Scores:

Pred. No.: 2,31e-34 Length: 204485
 Score: 563.00 Matches: 115
 Percent Similarity: 54.52% Conservative: 72
 Best Local Similarity: 33.53% Mismatches: 128
 Query Match: 26.36% Indels: 28
 DB: 2 Gaps: 7

US-09-208-629F-3 (1-407) x AC114962 (1-204485)

Qy 49 ThrLeuThrIleLysSerPheAsnGlyGlyProGlnAsnThrPheGluGluPheProLeu 68
 Db 73801 ACCCTGTCTTCCTTCTTGACAGGAAACCAATAGACCTCTAAAGGAAGAGCCTTATT 73860
 Qy 69 SerAspIleGluGlyTrp-----ThrGlyAlaThrThrIleLysAlaGluCys 85
 Db 73861 GGTAAGGTGTAGGCACATCCACGTCACCTGGAAGAGGTTACAGTTGAA----- 73911
 Qy 86 ProGluAspSerIleSerThrLeuHisValAsnAsnAlaThrIleGlyTyrLeuArgSer 105
 Db 73912 -----ACAGTCTTTCTGTGTGATGAGTTTCTGTCATCTGCTCCTCAGCTGGA 73956
 Qy 106 SerLeuSerThrGlnValIleProAlaIleTyrIleLeuLeuPheValValGlyValpro 125
 Db 73957 AAACGTACCATCTGCTTCCTTCCAAATGCTACCAATGTTGTTGGGGTTGCCA 74016
 Qy 126 SerAsnIleValThrLeuTrpLysLeuSerLeuArgThr---LysSerIleSerLeuVal 144
 Db 74017 AGTAAGCGCATGGCCCTGTGGTCTTTCTTTCCGAACCTAAGAAAGACCCCTGCTGTG 74076
 Qy 145 IlePheHisThrAsnLeuAlaIleAlaAspLeuPheCysValThrLeuProPheLys 164
 Db 74077 ATTTACATGGGCAATCTGGCCTTGCTGCTACCTCTCTCTGTCATCTGTTCCCTTGAAG 74136

Qy 165 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValMetCysArgIleThr 184
 Db 74137 ATTGCTATCACATACATGGCAACACCTGATTTATGGGAAGCTCTTTGTAATGCTT 74196
 Qy 185 ThrValValPheTyrGlyAsnMetTyrCysAlaIleLeuIleLeuThrCysMetGlyIle 204
 Db 74197 ATTGGCTTTTCTATGGCAACATGACTGCTTCATTTCTTTCATGACCTGCTCAGTGTG 74256
 Qy 205 AsnArgTyrLeuAlaThrAlaHisProPheThrTyrGlnLysLeuProLysArgSerPhe 224
 Db 74257 CAGAGTATTGGTTCATCTGAAACCCCTGAGGCTCCTCAGG---AAGAGGCAACATT 74313
 Qy 225 SerLeuLeuMetCysGlyIleValTrpValMetValPheLeuTyrMetLeuProPheVal 244
 Db 74314 GCCATTGGCATCTCCCTGGCAATATGGCTGCTGATTTCTCTGCTCACCCTCTTGTAT 74373
 Qy 245 IleLeuLysGlnGluTyrHisLeuValHisSerGluIleThrThrCysHisAspValVal 264
 Db 74374 GTCGTGAAGCAGACCATCTTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 74433
 Qy 265 AspAlaCysGluSerProSerSerPhe-----ArgPheTyrTyrPheValSer 280
 Db 74434 -----CCTGAGCAGCTCTTGGTGGGAGACATGTTCAATTACTTCTCTCT 74478
 Qy 281 LeuAlaPhePheGlyPheLeuIleProPheValIleIleIlePheCysTyrThrThrLeu 300
 Db 74479 CTGGCCATTGGGTCTTTCTGTTCCAGCCTTCTCCACAGCCTCTGCTGCTGCTGCTGCT 74538
 Qy 301 IleHisLysLeuLysSer-----LysAspArgIleTrpLeuGlyTyr 314
 Db 74539 ATCAGAATCTGCGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 74598
 Qy 315 IleLysAlaValLeuLeuLeuIlePheThrIleCysPheAlaProThrAsnIle 334
 Db 74599 ATCAAACTCATGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 74658
 Qy 335 IleLeuValIleHisAlaAsnTyrTyrTyrHisAsnThrAspSerLeuTyrPheMet 354
 Db 74659 CTGCTTGTGTCATTTCTGATTAGAGCCAGGCGCAGACCATGCTCATGCTGCTGCTGCT 74718
 Qy 355 TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPhe 374
 Db 74719 TACATTGTAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 74778
 Qy 375 ValMetSer 377
 Db 74779 TTTGTTTCA 74787

RESULT 38

HSU36753 1124 bp DNA linear PRI 28-FEB-1996
 LOCUS Human protease-activated receptor 2 gene, exon 2 and partial cds.
 DEFINITION U36753
 ACCESSION U36753.1 GI:1208539
 VERSION
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1124)
 AUTHORS Kahn,M.L. and Coughlin,S.R.
 TITLE Direct Submission
 JOURNAL Submitted (19-SEP-1995) Mark L. Kahn, CVRI/Medicine, UCSF, 3rd and
 Parnassus, San Francisco, CA 94143-0524, USA
 FEATURES
 Location/Qualifiers
 1. 1124
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /map="5q13"
 /note="Human genomic DNA in P1 bacteriophage library"
 <3..1112
 /codon_start=1
 CDS

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 88955)

Worley, K.C.

Direct Submission

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 88955)

Worley, K.C.

Direct Submission

Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GZLT

Center clone name: CH230-498A19

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 35763 bases at least Q40

Consensus quality: 40497 bases at least Q30

Consensus quality: 44674 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 44 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1023: contig of 1023 bp in length

* 1024 1123: gap of unknown length

* 1124 2158: contig of 1035 bp in length

* 2159 2258: gap of unknown length

* 2259 3396: contig of 1138 bp in length

* 3397 3496: gap of unknown length

* 3497 4759: contig of 1263 bp in length

* 4760 4859: gap of unknown length

* 4860 6321: contig of 1462 bp in length

* 6322 7471: gap of unknown length

* 7472 7571: gap of unknown length

* 7572 8781: contig of 1210 bp in length

* 8782 8881: gap of unknown length

* 8882 10238: contig of 1357 bp in length

* 10239 10339: gap of unknown length

* 10339 11514: contig of 1176 bp in length

* 11515 11614: gap of unknown length

* 11615 13207: contig of 1593 bp in length

* 13208 13307: gap of unknown length

* 13308 14714: contig of 1407 bp in length

* 14715 14814: gap of unknown length

* 14815 16170: contig of 1356 bp in length

* 16171 16270: gap of unknown length

* 16271 17526: contig of 1256 bp in length

* 17527 17626: gap of unknown length

* 17627 19027: contig of 1401 bp in length

* 19028 19127: gap of unknown length

* 19128 20258: contig of 1131 bp in length

* 20259 20358: gap of unknown length

* 20359 21562: contig of 1204 bp in length

* 21563 21662: gap of unknown length

* 21663 23425: contig of 1763 bp in length

* 23426 23525: gap of unknown length

* 23526 24837: contig of 1312 bp in length

* 24838

* 24937: gap of unknown length

* 24938 25946: contig of 1009 bp in length

* 25947 26046: gap of unknown length

* 26047 27583: contig of 1537 bp in length

* 27584 27683: gap of unknown length

* 27684 29343: contig of 1660 bp in length

* 29344 29443: gap of unknown length

* 29444 30864: contig of 1421 bp in length

* 30865 30964: gap of unknown length

* 30965 32341: contig of 1377 bp in length

* 32342 32441: gap of unknown length

* 32442 34592: contig of 2051 bp in length

* 34593 34937: gap of unknown length

* 34938 37027: contig of 2435 bp in length

* 37028 37127: gap of unknown length

* 37128 40023: contig of 2896 bp in length

* 40024 40123: gap of unknown length

* 40124 42440: contig of 2317 bp in length

* 42441 42540: gap of unknown length

* 42541 45193: contig of 2653 bp in length

* 45194 45293: gap of unknown length

* 45294 47982: contig of 2689 bp in length

* 47983 48082: gap of unknown length

* 48083 49458: contig of 1376 bp in length

* 49459 49558: gap of unknown length

* 49559 51558: contig of 2000 bp in length

* 51560 51658: gap of unknown length

* 51659 53031: contig of 1373 bp in length

* 53032 53131: gap of unknown length

* 53132 55703: contig of 2572 bp in length

* 55704 55803: gap of unknown length

* 55804 57502: contig of 1699 bp in length

* 57503 57602: gap of unknown length

* 57603 59938: contig of 2336 bp in length

* 59939 60038: gap of unknown length

* 60039 62314: contig of 2276 bp in length

* 62315 62414: gap of unknown length

* 62415 65536: contig of 3122 bp in length

* 65537 65636: gap of unknown length

* 65637 68901: contig of 3265 bp in length

* 68902 69001: gap of unknown length

* 69002 71415: contig of 2413 bp in length

* 71416 71514: gap of unknown length

* 71515 74439: contig of 2925 bp in length

* 74440 74539: gap of unknown length

* 74540 78489: contig of 3950 bp in length

* 78490 78589: gap of unknown length

* 78590 82200: contig of 3610 bp in length

* 82201 82299: gap of unknown length

* 82300 85743: contig of 3444 bp in length

* 85744 85843: gap of unknown length

* 85844 88955: contig of 3112 bp in length.

FEATURES

Location/Qualifiers

source

1..88955

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

/clone="CH230-498A19"

BASE COUNT 22306 a 20029 c 19951 g 21690 t 4379 others

ORIGIN

Alignment Scores:

Pred. No.: 1..75e-34

Score: 559.50

Percent Similarity: 56.88%

Best Local Similarity: 36.25%

Query Match: 26.19%

DB: 2

US-09-208-629F-3 (1-407) x AC128225 (1-88955)

Length: 88955

Matches: 116

Conservative: 66

Mismatch: 121

Indels: 18

Gaps: 6

Qy 95 ValAsnAlaThrLeGlyTyrLeuArgSerLeuSerThrGlnValProAla 114

Db 33815 GTTGATGATTCCTCTGCATCCGTCCTCACCGGAAGCTGACCGCTTTCTCCGGTC 33756

QY	115	IleTyrIleLeuLeuPheValValGlyValProSerAsnIleValThrLeuTyrPlyLeu	134
Db	33755	ATCTACATCATGTTCTTTGTAATAGTGTGGCCAGTAATGATGGCCCTCTGGGTCTTC	33696
QY	135	SerLeuArgThr--LysSerIleSerLeuValIlePheHisThrAsnLeuAlaIleAla	153
Db	33695	TTCTCCGACGAGAGAGAGACCCCTGCTGTGATTTATCATGGCCACACTGGCCTTGCGCA	33636
QY	154	AspLeuLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsn	173
Db	33635	GACCTCCTCTCTGTCATCTGGTTCCTCCCTGAAGATCTCTACCACTCCATGGCAACGAC	33576
QY	174	TrpValPheGlyGluValMetCysArgIleThrThrValValPheTyrGlyAsnMetTyr	193
Db	33575	TGGACCTATGGGATGCGCTCTGCAAGGTGCTCATTTGGCTTTTCTACGGCAATATGTAC	33516
QY	194	CysAlaIleLeuLeuThrCysMetGlyIleAsnArgTyrLeuAlaThrAlaHisPro	213
Db	33515	TGCTCATCTCTTTATGAACATGCTCTACGGTGCAGAGGTACTGGGTATCGTCAACCCC	33456
QY	214	PheThrTyrGlnLysLeuProLysArgSerPheSerLeuLeuMetCysGlyIle----	231
Db	33455	ATGGGACAC-----TCCAGGAGAGGGGCAACATCGCTGTGTGGCGTCTCCCTG	33408
QY	232	---ValTrpValMetValPheLeuTyrMetLeuProPheValIleLeuLysGlnGluTyr	250
Db	33407	GGCATCTGGCTCTGATTTTCTGGTCACCATCCCTCTGTAGTTCATGAGCAGAACATC	33348
QY	251	HisLeuValHisSerGluIleThrThrCysHisAspValValAspAlaCysGluSerPro	270
Db	33347	TACATTCCAGCCTTGACATACCACTGTTCACGACGTGCTGCCCGAG---GAGGTCTCTG	33291
QY	271	SerSerPheArgPheTyrThrPheValSerLeuAlaPhePheGlyPheLeuIleProPhe	290
Db	33290	GTGGGGGACATGTTTCAGTTACTTCTCTCCCTGGCCATGGAGTCTTTCTGTCCACGCC	33231
QY	291	ValIleIlePheCysTyrThrThrLeuIleHisLysLeuLysSer-----	306
Db	33230	CTCTTTACTCGCTCTGCCTACGTGCTCATGATCAAAACGCTCCGTCTCCGCCATGGAC	33171
QY	307	-----LysAspArgIleTrpLeuGlyTyrIleLysAlaValLeuLeuValIle	324
Db	33170	GAGCACTCGGAGAGAAAGGCGAGGGGTATCGCCTCATCATCACGTCTCTCCATG	33111
QY	325	PheThrIleCysPheAlaProThrAsnIleIleValIleHisHisAlaAsnTyrTyr	344
Db	33110	TACTTCATCTGCTTGCCTCCAGAACGTGCTGCTGTCGTGCAATATTCTCTCATCAA	33051
QY	345	TyrHisAsnThrAspSerLeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeu	364
Db	33050	AGCCAGAGGCAGAGCCACGTCTACGCCCTCTACCTCGTCGCCCTCTGCTCTCCACCTC	32991
QY	365	AsnSerCysLeuAspProPheLeuTyrPheValMetSerLysValValAspGlnLeuAsn	384
Db	32990	AACAGCTGATAGACCCCTTTGCTACTACTTTGTTTCGAAA-----GATTT-CAGGGA	32938
QY	385	Pro**SerAlaMetAlaArgProLeu**ArgProArgArgAspIleTrpGluAspIle	404
Db	32937	CNAGGCCAGAAAGCGCTCCTCTCGCGAAGCGGTCCGCAACCGTGAACGCATGCGAGATATC	32878
RESULT	40		
AR012640			
LOCUS		1414 bp	DNA linear PAT 04-DEC-1998
DEFINITION		Sequence 62 from patent US 5763575.	
ACCESSION		AR012640	
VERSION		AR012640.1	GI:3970630
KEYWORDS			
SOURCE		Unknown.	
ORGANISM		Unknown.	
REFERENCE		1 (bases 1 to 1414)	
AUTHORS		Sundelin, J. and Scarborough, R.M.	

```

Qy 297 TyrThrThrLeuIleHisLysLeuLysSer-----LysAspArgIle 310
Db 833 TATCTGCTGATGATCAGAAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAGAAA 892
Qy 311 TrpLeuGlyTyrIleLysAlaValLeuLeuIleLeuValIlePheThrIleCysPheAla 330
Db 893 AGGAAGAGGGCCATCAAACTCATGTCTCTGGGCATGTACCTGATCTGCTTCACT 952
Qy 331 ProThrAsnIleIleLeuValIleHisHisAlaAsnTyrTyrTyrHisAsnThrAspSer 350
Db 953 CCTAGTAACCTTCTGCTGTGGTGCATTATTTCTGATTAAGAGCCAGGCCAGAGCCAT 1012
Qy 351 LeuTyrPheMetTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 370
Db 1013 GTCTATGCCCTGTACATGTAGCCCTCTGCTCTCTACCCCTTACAGCTGCATCGACCCC 1072
Qy 371 PheLeuTyrPheValMetSer 377
Db 1073 TTGCTCTATTACTTTGTTCA 1093

```

Search completed: June 29, 2003, 09:40:41
Job time : 2691.22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:58:31 ; Search time 51.4368 Seconds
(without alignments)
1630.373 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLYTXQHPVAGSDIKMKIL.....AMARPLXPRRDIEDIHAW 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mmc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriapi.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	580.5	27.2	399	11 Q8R3I1	Q8R3I1 mus musculus
2	429	20.1	374	13 Q57466	Q57466 meleagris g
3	418	19.6	357	13 Q9DE05	Q9DE05 raja erinac
4	388.5	18.2	361	13 Q9OX57	Q9OX57 xenopus lae
5	370.5	17.3	309	11 Q8R528	Q8R528 mus musculus
6	369.5	17.3	337	4 Q96P68	Q96P68 homo sapien
7	344.5	16.1	298	11 Q8VE54	Q8VE54 mus musculus
8	342	16.0	377	13 Q98U14	Q98U14 brachydanio
9	338	15.8	337	4 Q75819	Q75819 homo sapien
10	336	15.7	330	4 Q9BXA5	Q9BXA5 homo sapien
11	336	15.7	334	4 Q8TDQ8	Q8TDQ8 homo sapien
12	336	15.7	359	11 Q99MT7	Q99MT7 mus musculus
13	335	15.7	359	11 Q99PP3	Q99PP3 cavia porce
14	333.5	15.6	358	4 Q9BY21	Q9BY21 homo sapien
15	333.5	15.6	358	4 Q96DZ8	Q96DZ8 homo sapien
16	330	15.4	351	12 Q9DHV5	Q9DHV5 yaba-like d

ALIGNMENTS

RESULT 1

```

Q8R3I1 ID Q8R3I1 PRELIMINARY; PRT; 399 AA.
AC Q8R3I1;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Coagulation factor II (thrombin) receptor-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025432; AAH25432.1; -
KW Receptor.
SQ SEQUENCE 399 AA; 44679 MW; 807C79464AB9B3EF CRC64;

```

Query Match 27.2%; Score 580.5; DB 11; Length 399;
Best Local Similarity 37.6%; Pred. No. 1.5e-41;
Matches 111; Conservative 64; Mismatches 103; Indels 17; Gaps 5;

```

QY 95 VNNATIGYLRSSLSQVIPAIVILLFVGVGPSNIVTLWKLURT-KSISLVFHTNLATA 153
Db 63 IDEFSASILTGLTITVFLPVVYIIIVFVIGLPSNGMALIFLFRTKKHPAVIYMANLALA 122
QY 154 DLLFCVTLTPFKIAYHLNGNNVFGVWCRTITVVFYGNMYCAIILITCNGINRYLATAHP 213
Db 123 DLLSVIWFPLATAYHLHGNWYVYGEALCKVLGFFYGNMYCSILFMTCLSVQRYWVNP 182
QY 214 FTYQKLPKESFSLMGIVWVFLYMLFVLKQBYHLVHSEITTCNDVVDACESSPSF 273
Db 183 MGHPR-KKANIAGVSLAIWLILFVTIPLYVMKQTIYIPALNITTCNDVLPD-EVLVD 240
QY 274 RYFYFVSLAFFGLPFPFVILIIICYITLI-----HKLKSKDRILWGIYKAVILLV 323
Db 241 MFNYFLSLAIGVFLFPAILTASAYVLMITLRSAMDEHSEKKRQRA----IRLIITVLA 296

```

```
QY 324 IFTICAPTNIILVIHANYHYHTDSLYFMYLIALCLGSLNSCLDPLFYFVMSK 378
DB 297 MYFICAPSMLLVHFLIKTORQSHVYALYLVALCLSTLNSCIDPFYFVFSK 351

RESULT 2
O57466 PRELIMINARY; PRT; 374 AA.
AC O57466;
DT 01-JUN-1998 (TEMBLrel. 06, Created)
DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE G protein coupled P2Y nucleotide receptor.
GN TP2Y.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=98086419; PubMed=9415702;
RA Boyer J.L., Waldo G.L., Harden T.K.;
RT "Molecular cloning and expression of an avian G protein-coupled P2Y
  receptor.";
RL Mol. Pharmacol. 52:928-934(1997).
DR EMBL; AF031897; AAC60339.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 374 AA; 42594 MW; 849C465722BDD02B CRC64;

Query Match 20.1%; Score 429; DB 13; Length 374;
Best Local Similarity 31.5%; Pred. No. 1.2e-28;
Matches 102; Conservative 56; Mismatches 118; Indels 48; Gaps 10;

QY 74 WTGATTIKAE--CPEDSISTLHVNATIGYLRSSLSSTQVIPAIIYLLFVGVPSNIVTL 131
DB 20 WLGNITAAAEAKC-----VFNEERKFI-----LLPISYGVFVVGGLPLNSWAM 63

QY 132 WKLSRTRKS-ISLVIFHTNLAIALDLFCVTLPFKIAHYHLNGNNVGEVWCRITTVFVG 190
DB 64 WIFVSRMRPNWNTTTFNFMNLAISDTLYVFSLSPLTVVYYADRNWPFQKVFCKIVRELFYA 123

QY 191 NMVCAIILTCMGINRYLATAHPTQKLPKRFSLLMCGIWMVFLYMLRPVILKQEV 250
DB 124 NLVSSILFUTCSVHRVMGCHPISLKWVTKHARLICVGLWLVVTCILPFLNIIFVTS 183

QY 251 HLHVSEITTCDDVDVADACESPSFRFY--YFVSLAFGLFIPFVIFVCTTLIHKL----- 304
DB 184 SKDNS--TLCHDTT---KPEEDHDVHYSSSIMALLFGIPFLVIVVCYCLMAKRLCKRS 237

QY 305 -----KSKDRILGYIKAVLLIIVFTICFAPTNIILVHANYHY-----HNTDSL 351
DB 238 FPSPSRPVPSYKXR-----SIKMIIVLVTPAICFVPFHTRTLTYTSRYFQADQCLNTII 293

QY 352 YFMVLIACLGSLNSCLDPLFYFV 375
DB 294 NFKYKTRPLASINSLDPLIYFM 317

RESULT 3
Q9DE05 PRELIMINARY; PRT; 357 AA.
AC Q9DE05;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
```

```
DE P2Y receptor.
OS Raja erinacea (Little skate).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalia; Hymnosquala; Pristiogaster; Batoidae;
OC Rajiformes; Rajidae; Raja.
OX NCBI_TaxID=7782;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20459151; PubMed=10900200;
RA Drancoff J.A., O'Neill A.F., Franco A.M., Cai S.Y., Connolly G.C.,
RA Ballatori N., Boyer J.L., Nathanson M.H.;
RT "A primitive ATP receptor from the little skate Raja erinacea.";
RL J. Biol. Chem. 275:30701-30706(2000).
DR EMBL; AF242850; AAG42684.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 357 AA; 41239 MW; 14604BE15DCBD41 CRC64;

Query Match 19.6%; Score 418; DB 13; Length 357;
Best Local Similarity 29.4%; Pred. No. 1e-27;
Matches 96; Conservative 62; Mismatches 131; Indels 38; Gaps 9;

QY 62 TFEF--PLSDIEGTGATTIKAECPDSISTLHVNATIGYLRSSLSSTQVIPAIIYLL 119
DB 8 TMEEVFANLSDLTGGLNKTC-----LNKGFOFYLYPLMIYIV 45

QY 120 FVGVPSNIVTLWKLSRTRKS-ISLVIFHTNLAIALDLFCVTLPFKIAHYHLNGNNVGE 178
DB 46 FVTGFIGNSVALLMFIHMRPWSITIIYFNLVLADLFVFSLSPLLIIFYFNKTDWIFGE 105

QY 179 VMCRITTVFVGNMYCAIILTCMGINRYLATAHPTQKLPKRFSLLMCGIWMVFL 238
DB 106 LLCKLXRFHFVNLVGSILFTICISVHRVTVGVHPMKSLGRLLKKSATIVCVVITVMA 165

QY 239 YMLPFVILKQEVHLVHSEITTCDDVDVADACESPSFRFYFVSLAFGLFIPFVIFVCT 298
DB 166 GISPLIYFSRT-GLRNNKNTCYDTTSK-ELLETY-FIYSMSTTFGFCIPATILVCYG 222

QY 299 TLIHKLKSKDR---IWLGVIKAVLLIIVFTICFAPTNIILVH-HANYHYHTDS---- 350
DB 223 FIVKALISNDMKTPLRGKSVRLIIVLAVFAISYLPFHVMKNLNDQSLRYOGLDTCWN 282

QY 351 --LYFMYLIALCLGSLNSCLDPLFYFV 375
DB 283 RVVATYQVTRGLASINSLDPLIYFL 309

RESULT 4
Q90X57 PRELIMINARY; PRT; 361 AA.
AC Q90X57;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE P2Y1 nucleotide receptor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheng A.W., Tsim K.W.;
RT "Cloning of Xenopus P2Y1 Receptor.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF432354; AAL27614.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
```


DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 361 AA; 41002 MW; 55B2D605F5B57FED CRC64;
Query Match 18.2%; Score 388.5; DB 13; Length 361;
Best Local Similarity 29.2%; Pred. No. 3.3e-25;
Matches 91; Conservative 67; Mismatches 115; Indels 39; Gaps 10;
QY 91 STLHVNATIGYLRSLSTQV-----IPAIYLLFVGVGPSNIVTLWKLKSLRK---SI 141-
DB 15 STLLASGSGAGNTKCSLTKFTGQFYLPFAVIVVCITGFIGNSVAIWMFIFHKWPSI 74
QY 142 SLVIFHTNLAIADLLFCVTLFPKFIAYHLGNMNVFGEVMCRITTVFYGNMYCAILLFRC 201
DB 75 SVYMF--NLALADFLVLSLPALIFYFNKTDWIFGDALCKLQRFELFHNLYGSLFLFC 132
QY 202 MGINRYLATAHPTQYOKLPKRSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTC 261
DB 133 ISVHRYTGVVHPPLKSLGRLLKKNISYISALVWFIVIAIGISP-ILFFSGTGIRKNKTTICF 191
QY 262 DVVDACESPSPRRFYFVSL--AFFGFLIPFVILIFCYTTLIHKLKSKD-----RI 310
DB 192 DT-----SSDEYLSYFYISNCTTVFGCPFIIDILGCGILVVALYKDMNAPLRKSI 247
QY 311 WLGIYKAVILLIIVIFITCFAPTNIILVIH-HANYYYH-----TDSLIFYMYLIALCLGS 363
DB 248 YL-----VIIVLTFAVSILPFHVMKNLRLARLDFQSPMCNFDNRVYATVQVTRGLAS 302
QY 364 LNSCLDPFLYFV 375
DB 303 LNSCVDPILYFL 314
RESULT 5
Q8R528 PRELIMINARY; PRT; 309 AA.
AC Q8R528;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cysteinyll leukotriene 2 receptor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishii S.;
RL "Mouse CysLT2 Gene."; Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB058930; BAB86881.1; -.
KW Receptor.
SQ SEQUENCE 309 AA; 35226 MW; 96FACC6B8AF6974 CRC64;
Query Match 17.3%; Score 370.5; DB 11; Length 309;
Best Local Similarity 29.3%; Pred. No. 9.8e-24;
Matches 85; Conservative 62; Mismatches 118; Indels 25; Gaps 8;
QY 97 NATIGYRLSLSQVIPAIIYLLFVGVGPSNIVTLWKLKSLRT--KSLVIFHTNLAIAD 154
DB 14 NCTI---ENFKEFYPIIIVLIFFWALGNGFSIV-VFLQTCCKSTSVNVFNLNATSD 68
QY 155 LLECVTLFPKFIAYHLGNMNVFGEVMCRITTVFYGNMYCAILLTCMGINRYLATAHFP 214
DB 69 FLFISLTPFPADYFRGNSWIFGDLACRVMSYSLYVNMVTSIYFLVLSVVRFLATVHPF 128
QY 215 TYQKLPKRSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTCDDVDACESPSPRR 274
DB 129 RMFHVTSVRSAILCGIHWFIWASSALLVNGQE---EKDNIISCLEL-----SQPKF 180
QY 275 FYYFVS--LAFFGFLIPFVILIFCYTTLIHKLKSKDRIWLG-----YIKAVLLILVIF 325

DB 181 SLLIMNHIAVAVGFLPLFTLTCTICLLIIRLLKAEIPESGPRAAHRKALTTIVIAMITF 240
QY 326 TICFAPTNIILVIHANYYYHNT-DSLIFYMYLIALCLGSLNSCLDPFLYF 374
DB 241 LLCFLFPHYHALRTLHLVTDKOSCGDVLHKATVITLTMAAANSCFNPFLLY 290
RESULT 6
Q96P68 PRELIMINARY; PRT; 337 AA.
AC Q96P68;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE G protein-coupled receptor (Putative G-protein coupled receptor).
GN GPR80 OR GPCR
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21458557; PubMed=11574155;
RA Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vanti W.B., Arkhitko O., Lewis T., Evans J.F., George S.R., O'Dowd B.F.;
RT "Discovery and mapping of ten novel G protein-coupled receptor genes";
RL Gene: 275:83-91(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT "Identification of G protein-coupled receptor genes from the human genome sequence";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF411109; AAL26480.1; -.
DR EMBL; AB083598; BAB89311.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 337 AA; 38251 MW; 6814EA0044756CE6 CRC64;
Query Match 17.3%; Score 369.5; DB 4; Length 337;
Best Local Similarity 29.6%; Pred. No. 1.3e-23;
Matches 87; Conservative 58; Mismatches 108; Indels 41; Gaps 8;
QY 107 LSTQVIPAIYLLFVGVGPSNIVTL-----WKLSLRTKSLVIFHTNLAIADLLF 157
DB 31 LKWHYLPVIYGIIFLVGFGNAVISTVIFKMRPWK-----STIIMLNLACTDLY 82
QY 158 CVTLFPKFIAYHLGNMNVFGEVMCRITTVFYGNMYCAILLTCMGINRYLATAHPTQY 217
DB 83 LTSLPLFIHYASGENWIPGDFMCKFIRFSPHNFYSSILFUTCSFYCYVIIHPMSCF 142
QY 218 KLPKRSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTCDDVDACESPSPRRFY 277
DB 143 SIHKTRCAVAVCAVWIIISLAVIPMTFLTSTN--RTNRSACDLTSSDEL-NTIKWYN 199
QY 278 FVSLAFFGFLIPFVILIFCYTTLIHK-----LKSRIWILGYIKAVILLIIVFTI 327
DB 200 LI-LTATFCLPLVITVLCYTTIHTLTHGLQDSCCLKOKAR-----RLTILLLAFV 252
QY 328 CFAPTNIILVIHANYYYHNT-----DSLIFYMYLIALCLGSLNSCLDPFLYFVMS 377
DB 253 CFLPPIHLIRVIRIESRLSISCSISQNIHQHEAVIVSRPLAALNTFGLNLLYVWVS 306
RESULT 7
Q8VE54 PRELIMINARY; PRT; 298 AA.
AC Q8VE54;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to G protein-coupled receptor 23 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.,
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019743; AAI19743.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; UNKNOWN_1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 298 AA; 33708 MW; A09209FEABCE65F2 CRC64;

Query Match 16.1%; Score 344.5; DB 11; Length 298;
Best Local Similarity 30.3%; Pred. No. 1.6e-21;
Matches 76; Conservative 52; Mismatches 88; Indels 35; Gaps 7;

QY 145 IFTNLAIADLLFCVTLPPKIAVHLNNGNWNVFGVWMCRTITVVFYGNMYCAILLILCMGI 204
DB 4 IFTNLALSDLLFCVTLPPKIFNFN-RHWPFGDTLCKSGTAFNLINYGSMFLFCISV 62

QY 205 NRYLATAHPPTYQKLPKRSFSLMCGIWMVVMFLYMLPFVILKQYHLVHSEITTCDDVV 264
DB 63 DRELAIVYPPRSRTIRTRNSAIVCAGVILVLSGGISASLF-----STTVNNAT 113

QY 265 DACESPSSRFY--YFVSLAFF-----GFLPFIIVIFCYTTLIHLK 308
DB 114 TTCFEGSKRWKTYLSKITIFTEVVGFIPLNLVNSCSVWLRLKPKATLSIGQTNKK 173

QY 309 RIWLGIKAVLLILVITFCFAPTNILVIHANYHYHNTDSLY-----FMYIALCLGS 363
DB 174 KV-----LKMITHVAVFVGVFVNSVFLYALVRSQAINTCLLERFAKIMVPIITCLAT 229

QY 364 LNSCLDPFIYF 374
DB 230 LNCCLDPFIY 240

RESULT 8
Q98U14 PRELIMINARY; PRT; 377 AA.
AC Q98U14;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Opioid receptor-like protein ZFOR3.
DE Opioid receptor-like protein ZFOR3.
OC Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez R.E., Gonzalez-Sarmiento R., Alvar F.;
RT "Cloning and characterization of ZFOR3, a new zebrafish opioid
receptor-like protein."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF285173; AAG60607.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.

KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 377 AA; 42306 MW; A8AAD07B345064E0 CRC64;

Query Match 16.0%; Score 342; DB 13; Length 377;
Best Local Similarity 27.2%; Pred. No. 3.2e-21;
Matches 86; Conservative 67; Mismatches 125; Indels 38; Gaps 8;

QY 86 PEDSISTLHVNATIGYLRSSLSLQ-----VIPAIYILLFVVGPSNIVTW 132
DB 19 PEKCLPNETWQSAVSDIYNSSLNESWTTEQEA MSPLIITAVYVGVGLVGNCLVWY 78

QY 133 KLSLETK---SISLVIFHTNLAIADLLFCVTLPPKIAVHLNNGNWNVFGVWMCRTITVVFY 189
DB 79 VIIIRHTKMTATNIYIF--NLRVADALVTTTFQSTDYLL-NSWPFGEVGVCKVFISIDY 135

QY 190 GNMVCAILLITCMGINRYLATAHPPTYQKLPKRSFSLMCGIWMVVMFLYMLPFVILKQE 249
DB 136 YNMFTSIFTLTMSVDRVAVCHPVKALDFRTPMKAKIINILILWSSAAGIPAMVLGST 195

QY 250 YHLVHSEITTCDDVVDACESPSSRFY--YFVSLAFFGLFPIIVIFCYTTLIHLKLS 306
DB 196 Q--TNGGTEC-----ALQFPDPYVYWDLMKICVIFGFVAPLIIITVCYILMVLRLKS 248

QY 307 -----KDRIMLGIKAVLLILVITFCFAPTNILVIHANYHYHNTDSLYPMYLI 357
DB 249 VRLLSGSRKDRNLRRITRLVLVAVVAVVWCWTPHIFILVKALSHGVPEITTSVMAAYFF 308

QY 358 ALCLGSLNSCLDPPELY 373
DB 309 CVALGYTNSLNPILY 324

RESULT 9
Q75819 PRELIMINARY; PRT; 337 AA.
AC Q75819;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE T cell death associated protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98317948; PubMed=9655242;
RA Kyaw H., Zeng Z., Su K., Fan P., Shell B.K., Carter K.C., Li Y.;
RT "Cloning, characterization, and mapping of human homolog of mouse T-
cell death-associated gene."
RL DNA Cell Biol. 17:493-500(1998).
DR EMBL; U95218; AAC31794.1; -.
DR HSSP; F34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR01649; PSYCHOSINER.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; UNKNOWN_1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
SQ SEQUENCE 337 AA; 39339 MW; 05D4739457B079A2 CRC64;

Query Match 15.8%; Score 338; DB 4; Length 337;
Best Local Similarity 28.4%; Pred. No. 6.3e-21;
Matches 94; Conservative 63; Mismatches 118; Indels 56; Gaps 12;

QY 97 NATIGYLRSSLSQTVIPAIYILLFVVGPSNIVTWLKLRLTKSIS-LVIFHTNLAIADL 155
DB 2 NSTCIEEQHDLHYLFPVIYIFVIIVIPANIGSLCVSLQPKSELGIYLFSLSLSDL 61

QY 156 LFCVTLPPKIAVHLNNGNWNVFGVWMCRTITVVFYGNMYCAILLITCMGINRYLATAHPPT 215
DB 62 LYALTPLPLWIDYTNKNDNWTFSFALCKGSAFLMYMKFYSSSTAFLTCIAVDRLAVVYPLK 121

```
QY 216 YOKLPKRSFSLMCGIWMVFLYMLPFVILKQEXHLVH-----SEITTCDDVVDACS 269
Db 122 FFPLRTRIALMVSLSIWILETIFNA--VMLWEDETWEYCDAEKSNFTLCYD-----KY 174
QY 270 PSSFRFYFVSLAFF---GFLIPFVILIFC-----YTLIH-----KLKSKDRILWLYIKA 317
Db 175 PLE---KQOINLMFTCTGYAIPILICNKKVYQAVRHNNKATENKEKKRI-----IKL 227
QY 318 VLLILVIFTCFAPTNIILVI-----HANYYYHNTDS--LYFMYLIALCLGSLNSCLDP 370
Db 228 LVSITVFLVCLFPFPHVLLIRLILEHAVNFEDHSNGKRTYMYEITVALISLNCVADP 287
QY 371 FLYFVMSKVVQDQNPXSAMARPLXRRPRDIW 401
Db 288 ILYCFVTET-----GRYDMM 302

RESULT 10
Q9BXAS PRELIMINARY; PRT; 330 AA.
AC Q9BXAS;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE G-protein coupled receptor 91.
GN GPR91.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
SQ SEQUENCE FROM N.A.
RX MEDLINE=21172992; PubMed=11273702;
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in
the discovery of new g-protein coupled receptors.";
RL J. Mol. Biol. 307:799-813(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF348078; AAK29080.1; -.
DR HSSP; P34996; iDDb.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS02642; G-PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS02642; G-PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 330 AA; 38283 MW; 6D3C70E340866BAB CRC64;

Query Match 15.7%; Score 336; DB 4; Length 330;
Best Local Similarity 30.0%; Pred. No. 9.1e-21;
Matches 86; Conservative 61; Mismatches 118; Indels 22; Gaps 8;

QY 105 SLSSTQVIPAIIYLLFVGVGPSNIVTLWK--LSLRT-KSISLVIFHTNLAIADLLFCVTL 161
Db 15 AALEKYLSIFYGIEFVGVGLNTIVVGYIFSLKNWSSNIYLF--NLSVSDLAFLCTL 72
QY 162 PFKIAHLNKNWVGEVWCRTITVVFYGNMYCAIILTCMGINRYLATAHFTYQKLPK 221
Db 73 PMLIRSYANG-NWYIGDVLCSNRYVLANLYTSILFTFISIDRYLIILKYPFREHLLQK 131
QY 222 RSFSLMCGIWMVFLYMLPFVILKQEXHLVHSEITTCDDVVDACESPSSFRFYFVSL 281
Db 132 KEFALISLAIWLVLTLELLP--ILPLINPVTDTGTTCTNDP--ASSGDPNLYISMCL 187
QY 282 AFGFLIPFIIFCYTTLIHKLSKDRILWLG-----IKAVLLILVIFTICFAPTNI 335
Db 188 TLGLGLIPLFVCMCFYKIALFLKQNRQVATAPLEKPLNLVIMAVIFSFLVTFYHVM 247
QY 336 LVIIHAN-----YYHNTDSLIFYMILIALCLGSLNSCLDPFLYFVM 376
Db 248 RNVRIASRLGSKWQYQCTQVINSFYIVTRPLAFLNSVINPVFYELL 294

Query Match 15.7%; Score 336; DB 4; Length 330;
Best Local Similarity 30.0%; Pred. No. 9.1e-21;
Matches 86; Conservative 61; Mismatches 118; Indels 22; Gaps 8;

QY 105 SLSSTQVIPAIIYLLFVGVGPSNIVTLWK--LSLRT-KSISLVIFHTNLAIADLLFCVTL 161
Db 19 AALEKYLSIFYGIEFVGVGLNTIVVGYIFSLKNWSSNIYLF--NLSVSDLAFLCTL 76
QY 162 PFKIAHLNKNWVGEVWCRTITVVFYGNMYCAIILTCMGINRYLATAHFTYQKLPK 221
Db 77 PMLIRSYANG-NWYIGDVLCSNRYVLANLYTSILFTFISIDRYLIILKYPFREHLLQK 135
QY 222 RSFSLMCGIWMVFLYMLPFVILKQEXHLVHSEITTCDDVVDACESPSSFRFYFVSL 281
Db 136 KEFALISLAIWLVLTLELLP--ILPLINPVTDTGTTCTNDP--ASSGDPNLYISMCL 191
QY 282 AFGFLIPFIIFCYTTLIHKLSKDRILWLG-----IKAVLLILVIFTICFAPTNI 335
Db 192 TLGLGLIPLFVCMCFYKIALFLKQNRQVATAPLEKPLNLVIMAVIFSFLVTFYHVM 251
QY 336 LVIIHAN-----YYHNTDSLIFYMILIALCLGSLNSCLDPFLYFVM 376
Db 252 RNVRIASRLGSKWQYQCTQVINSFYIVTRPLAFLNSVINPVFYELL 298

RESULT 12
Q99MT7 PRELIMINARY; PRT; 359 AA.
AC Q99MT7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE G-protein coupled receptor GPR87.
GN GPR87.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
SQ SEQUENCE FROM N.A.
RX MEDLINE=21172992; PubMed=11273702;
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in
the discovery of new g-protein coupled receptors.";
RL J. Mol. Biol. 307:799-813(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF295366; AAK01866.1; -.
DR MGD; MGI:1934133; Gpr87.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
```

```
RESULT 11
Q8TDQ8 PRELIMINARY; PRT; 334 AA.
AC Q8TDQ8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE P2Y purinoceptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
SQ SEQUENCE FROM N.A.
RA Zhang W., Li N., Wan T., Cao X.;
RT "Human P2Y purinoceptor 1.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF247785; AAL95690.1; -.
SQ SEQUENCE 334 AA; 38697 MW; 33146E1AD87F0E81 CRC64;

Query Match 15.7%; Score 336; DB 4; Length 334;
Best Local Similarity 30.0%; Pred. No. 9.2e-21;
Matches 86; Conservative 61; Mismatches 118; Indels 22; Gaps 8;

QY 105 SLSSTQVIPAIIYLLFVGVGPSNIVTLWK--LSLRT-KSISLVIFHTNLAIADLLFCVTL 161
Db 19 AALEKYLSIFYGIEFVGVGLNTIVVGYIFSLKNWSSNIYLF--NLSVSDLAFLCTL 76
QY 162 PFKIAHLNKNWVGEVWCRTITVVFYGNMYCAIILTCMGINRYLATAHFTYQKLPK 221
Db 77 PMLIRSYANG-NWYIGDVLCSNRYVLANLYTSILFTFISIDRYLIILKYPFREHLLQK 135
QY 222 RSFSLMCGIWMVFLYMLPFVILKQEXHLVHSEITTCDDVVDACESPSSFRFYFVSL 281
Db 136 KEFALISLAIWLVLTLELLP--ILPLINPVTDTGTTCTNDP--ASSGDPNLYISMCL 191
QY 282 AFGFLIPFIIFCYTTLIHKLSKDRILWLG-----IKAVLLILVIFTICFAPTNI 335
Db 192 TLGLGLIPLFVCMCFYKIALFLKQNRQVATAPLEKPLNLVIMAVIFSFLVTFYHVM 251
QY 336 LVIIHAN-----YYHNTDSLIFYMILIALCLGSLNSCLDPFLYFVM 376
Db 252 RNVRIASRLGSKWQYQCTQVINSFYIVTRPLAFLNSVINPVFYELL 298

RESULT 12
Q99MT7 PRELIMINARY; PRT; 359 AA.
AC Q99MT7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE G-protein coupled receptor GPR87.
GN GPR87.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
SQ SEQUENCE FROM N.A.
RX MEDLINE=21172992; PubMed=11273702;
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in
the discovery of new g-protein coupled receptors.";
RL J. Mol. Biol. 307:799-813(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF295366; AAK01866.1; -.
DR MGD; MGI:1934133; Gpr87.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
```

```
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS01655; UDPGLUCOSER.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.
DR G-PROTEIN COUPLED RECEPTOR; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 359 AA; 41462 MW; 2EB313C273EFCD6A CRC64;

Query Match 15.7%; Score 336; DB 11; Length 359;
Best Local Similarity 24.5%; Pred. No. 1e-20;
Matches 90; Conservative 72; Mismatches 123; Indels 82; Gaps 12;

QY 46 AKETLIKFGNGGQNTFFEFPLSDIEGTGATTIKACPDSDISTLHVNNTATIGYLS 105
DB 2 AVENVNVSTFAGN-----ELYSQASHT-----ANSTSEGHGKNSL---LHN 39

QY 106 SLSTQVIPAIIYLLFVVGVPNSIVTLW-KLSLRTKSIISLVFHTNLAIADLLFCVTLPLPK 164
DB 40 KPTIILPLVLYIFVASILLNGLAIVFFHIRNKT-SFIFLKNIVVADLIMTLTFPFR 98

QY 165 IAVHLGNWVFGVCMCRITTVVFGNMYCAIILITCMGINRYLATAPFTYOKLPKRSE 224
DB 99 IVRDAGFGPWYFFILCRYTSULFYANMYTSIVFLGLISVDRVLKVKVPGDSRMYSTIF 158

QY 225 SLAMCGIWMVWFLYMLPFVIL-----KOEYH-----LVHSEITTCCHDVVDAC 267
DB 159 TKVLSVCVMVIMAILSLPNILNTNGOPTKENIHDCMKLSPLGAKWHMAVT---YVDS 214

QY 268 ESPSSFRFYFVSLAFEGFLIPFVILFCVTLIHLKSKDRWLW-----YIKAVLL 320
DB 215 L-----FVAVILVILGCIYAIISRYIHKSSQFISQSSQSKRKHNSIRV 257

QY 321 ILVIFTICFAPTNILVIHANYHYNTDS-----LYFMYIALCLGSLNSCLDPF 371
DB 258 VVAVFTCFPLPHLCRI-----PTFNSLDLDESAAHLYLYCKENTLFLSACNCLDPI 313

QY 372 LYFVMSK 378
DB 314 IYFFMCK 320

RESULT 13
Q9EPP3 PRELIMINARY; PRT; 359 AA.
AC Q9EPP3;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Angiotensin II type 1 receptor.
GN At1.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=GALL BLADDER;
RA Busmann A., Noack T., Zucht HD., Forsemann W.G., Meyer M.;
RT "Isolation of Angiotensin I from human hemofiltrate as a potent
RT modulator of KATP channels."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AJ301623; CAC21550.1; -.
DR HSP; P34996; 1DD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.
DR G-PROTEIN COUPLED RECEPTOR; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 359 AA; 41045 MW; D00F4C2C472DE5CC CRC64;

Query Match 15.7%; Score 335; DB 11; Length 359;
```

```
Best Local Similarity 26.8%; Pred. No. 1.2e-20;
Matches 88; Conservative 54; Mismatches 122; Indels 64; Gaps 8;

QY 81 IKAECPEDSISTLHVNNTATIGYLSLSSTQVIPAIIYLLFVVGVPNS-----IVTLWLKSL 136
DB 14 IQDDCPKDG-----RHSYIFVMPTDYSIIFVVGIFGNSLVIVYFYWKL 59

QY 137 RTKSISLVIFHTNLAIADLLFCVTLIPFKIAYHLGNWVFGVCMCRITTVVFGNMYCAI 196
DB 60 KTVIA---SVFLNLALADICELLTLPLWAVYTAMEYRWPFGNMYCKIASASVSFLYASV 116

QY 197 ILITCMGINRYLATAPFTYOKLPKRSLMCGIWMVWFLYMLPFVILKOEYHLVHSE 256
DB 117 FLITCLSIDRYLAIVHPMKSLRLRMLVAKVTCVIIWLMAGLASLPAVIHRNVFFIENTN 176

QY 257 ITTCHDVVDACESPSFRFY-----FVSLAFEGFLIPFVILFCVTLIHLK 305
DB 177 ITVC-----AFHYESQNSTLPIGLGLTKNLGFMFPFLIILTSYTLIWKALK 223

QY 306 SKDRWLW-----YIKAVLLIIVFTICFAPTNIL-----VIHANYHYNTDS 350
DB 224 KAYELQKNKPRNDDIFKIIIMAVLFFPFSWPHQIFTFDLVLQLGLIHDCKI---SDI 279

QY 351 LYFMYIALCLGSLNSCLDPFLYFVMSK 378
DB 280 VDTAMPITICAYFNCLNPLFYGFGLK 307

RESULT 14
Q9BY21 PRELIMINARY; PRT; 358 AA.
AC Q9BY21;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Orphan G protein-coupled receptor 87.
GN GPR87 OR GPR95.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21172992; PubMed=11273702;
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in
RT the discovery of new g-protein coupled receptors."
RL J. Mol. Biol. 307:799-813(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21458557; PubMed=11574155;
RA Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vanti W.B., Arkhitko O.,
RA Lewis T., Evans J.F., George S.R., O'Dowd B.F.;
RT "Discovery and mapping of ten novel G protein-coupled receptor
RT genes."
RL Gene 275:83-91(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF237763; AAK01858.1; -.
DR EMBL; AF411114; AAL26485.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PRINTS; PR01655; UDPGLUCOSER.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.
DR G-PROTEIN COUPLED RECEPTOR; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 358 AA; 41435 MW; FDF157295BE5D10F CRC64;

Query Match 15.6%; Score 333.5; DB 4; Length 358;
Best Local Similarity 25.5%; Pred. No. 1.6e-20;
Matches 86; Conservative 76; Mismatches 132; Indels 43; Gaps 11;
```

```

QY 62 TPEPPLSDIEGWTGATTIKACRPSDSTLHVNNATIGYLRSSLSSTQVIPAIIYLLFV 121
Db 6 TLAKLNNELHGOESHNSGRSGPKN-TTLH-----NEEDTIVLPVLYLIIFV 54
QY 122 VGVPSNIVTLW-KLSLRKTSISLVIHTNLAIADLLFCVTLPPKIAHLNGNNVFEVYM 180
Db 55 ASILLNGLAVWIFFHIRNKT-SFIFYLKNIIVADLIMTLTFPRIVHDAGFGPWYKFKIL 113
QY 181 CRITTVVFGNMYCAIILTCMGINRYLATAHPTTYQKLPKRFSLLMCGIIVVMVFLYM 240
Db 114 CRYTSVLFVANNMTYSIVFLGLSIDRYLKVVKPGDSRMYSITFTKLVSCVVMIVAVLS 173
QY 241 LPVILKQEHVHSEITTCDDVDAC---ESPSSFRFY-----YFVSLAFFGLIPFVII 293
Db 174 LPNII-----LTNQPT--EDNIHDCSKLSPGVKWHATAVTVVNSCLFVAVL---VIL 222
QY 294 IFCYTTLIHLKSKDRWLIG-----YIKAVLLILVIFTICFAPNTIILVIHANNYYH 346
Db 223 IGCYIAISRYIHKSQRQFISQSRKRKHQNSIRVVAVFTCTPLPVLHCRIPFTFSHLDR 282
QY 347 NTDS-----LYFMYLIALCLGSLNSCLDPLFLYFVMSK 378
Db 283 LLDESAQKILYCKEITLFLSACNVCLDPIIYFFMCR 319

RESULT 15
Q96JZ8 PRELIMINARY; PRT; 358 AA.
AC Q96JZ8;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE CDNA FLJ14878 fis, clone PLACE1003238, weakly similar to probable G
DE protein-coupled receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuna M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamanoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027784; BAB5366.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1.1;
DR PRINTS; PR01569; F2Y12PRNCPT.
DR PRINTS; PR01655; UDPGLUCOSER.
DR PROSITE; PS00237; G PROTEIN RECEPT_F1.1; UNKNOWN_1.
DR PROSITE; PS0262; G PROTEIN RECEPT_F2.1.
SQ SEQUENCE 358 AA; 41462 MW; 7822700C8E3CC7E9 CRC64;

Query Match 15.6%; Score 333.5; DB 4; Length 358;
Best Local Similarity 25.5%; Pred. No. 1.6e-20;
Matches 86; Conservative 76; Mismatches 132; Indels 43; Gaps 11;

QY 62 TPEPPLSDIEGWTGATTIKACRPSDSTLHVNNATIGYLRSSLSSTQVIPAIIYLLFV 121
Db 6 TLAKLNNELHGOESHNSGRSGPKN-TTLH-----NEEDTIVLPVLYLIIFV 54
QY 122 VGVPSNIVTLW-KLSLRKTSISLVIHTNLAIADLLFCVTLPPKIAHLNGNNVFEVYM 180
Db 55 ASILLNGLAVWIFFHIRNKT-SFIFYLKNIIVADLIMTLTFPRIVHDAGFGPWYKFKIL 113
QY 181 CRITTVVFGNMYCAIILTCMGINRYLATAHPTTYQKLPKRFSLLMCGIIVVMVFLYM 240

```

```

Db 114 CRYTSVLFVANNMTYSIVFLGLSIDRYLKVVKPGDSRMYNITFTKLVSCVVMIVAVLS 173
QY 241 LPVILKQEHVHSEITTCDDVDAC---ESPSSFRFY-----YFVSLAFFGLIPFVII 293
Db 174 LPNII-----LTNQPT--EDNIHDCSKLSPGVKWHATAVTVVNSCLFVAVL---VIL 222
QY 294 IFCYTTLIHLKSKDRWLIG-----YIKAVLLILVIFTICFAPNTIILVIHANNYYH 346
Db 223 IGCYIAISRYIHKSQRQFISQSRKRKHQNSIRVVAVFTCTPLPVLHCRIPFTFSHLDR 282
QY 347 NTDS-----LYFMYLIALCLGSLNSCLDPLFLYFVMSK 378
Db 283 LLDESAQKILYCKEITLFLSACNVCLDPIIYFFMCR 319

RESULT 16
Q9DHV5 PRELIMINARY; PRT; 351 AA.
AC Q9DHV5;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE 7L protein.
GN 7L.
OS Yaba-like disease virus (YLDV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus.
OX NCBI_TaxID=132475;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21176366; PubMed=11277691;
RA Lee H.J., Essani K., Smith G.L.;
RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus.";
RL Virology 281:170-192 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Thesis (2000), Sir William Dunn School of Pathology, University of
DR EMBL; AJ293568; CAC21245.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PRINTS; PR00237; GPCRHOPOPSN.
DR PROSITE; PS0262; G PROTEIN RECEPT_F1.2; 1.
SQ SEQUENCE 351 AA; 41074 MW; FDDA643301887004 CRC64;

Query Match 15.4%; Score 330; DB 12; Length 351;
Best Local Similarity 27.8%; Pred. No. 3.2e-20;
Matches 82; Conservative 73; Mismatches 106; Indels 34; Gaps 13;

QY 91 STLHVNNATIGYLRSSLSSTQVIPAIIYLLFVGVPSNIVTLKLSLRKTSISLV--IPHT 148
Db 39 NVIYDNKNTI-----ISFVYCMFLIFGLLNLIVLVLT-CKLSLTVDIFLL 86
QY 149 NLAIADLLFCVTLPPKIAHLNGNNVFEVYCRITTVVFGNMYCAIILTCMGINRYL 208
Db 87 NLAIADLLFVMSFPFQIHYQL--SQMIFGNFCKIVSGLYIGFSGMFEVTVMSVYRI 144
QY 209 ATAHPTTYQ-KLPKRFSLLMCGIIVVMVFLMPLPFVILKQEHVHSEITTCDDVDAC 267
Db 145 SIHV-ITVSLKIKTVKIGYISLLLIWIISIVLTTPLVVYQVEK--HDTLCY----AF 197
QY 268 ESPSSFRFYFVS--LAPFGFLIPFVIIIFCVTTLIHLKLS-KDRIMLGVIKAVLLIIV 324
Db 198 YNNKTFIWLFINFENIIGMLIPFVILLFCYVKILMQLCKCKNKKIKAIRLVLIIFV 257
QY 325 FTICFAPNTIILVI--HHANNYYHTDSL-----YFMYLIALCLGSLNSCLDPLFLY 373
Db 258 NIIFWIPFNVLFTSMHSLFEGCKTFKITYALYVTEI-ISSSHCCINPLIY 311

```

RESULT 17

Q9PVY7 ID Q9PVY7 PRELIMINARY; PRT; 359 AA.
 AC Q9PVY7
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Angiotensin receptor.
 OS Anguilla anguilla (European freshwater eel).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
 OC Anguilla
 NCBI_TaxID=7936;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tran Van Chui M., Dolphin C.T., Barker S., Clark A.J., Vinson G.P.;
 RT "Molecular cloning and characterization of the cDNA encoding the
 RT angiotensin II receptor of european eel Anguilla anguilla."
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AJ005132; CAB40835.1; -;
 DR HSSP; P34996; 1DDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR000130; Zn_WTptdase.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPS.
 DR PROSITE; PS00237; G-PROTEIN RECEPT_F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPT_F1_2; 1.
 DR PROSITE; PS00290; Ig_MHC; UNKNOWN_1.
 DR PROSITE; PS00142; ZINC-PROTEASE; UNKNOWN_1.
 KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
 SQ SEQUENCE 359 AA; 39997 MW; 053726AEA827C271 CRC64;
 Query Match 15.4%; Score 328; DB 13; Length 359;
 Best Local Similarity 25.7%; Pred. No. 4.8e-20;
 Matches 82; Conservative 60; Mismatches 109; Indels 68; Gaps 9;
 QY 93 LHVNNATIGYRLSLSQVPAIYILLFVVGVPSPN--IVLWKLSLRKTSISLIVFHTNL 150
 DB 12 IHITCNTSG--RHSYIYTLIPVYGCNFGVIGVGNMVAIVCYMKLTIVA-NIFVLNL 68
 QY 151 AIADLFLFCVTLPKIAVHLNGNNWVGEVNCRTTTFVFGNMYCAILLITCMGINRYLAT 210
 DB 69 AVSDLTFLITLPMWATFTAMGYNWPFGGLCKASAGLITFNLYTSIFLTSLSDYLAI 128
 QY 211 AHPFTYOKLPKRSFSLMCGIWMVFLYMLPFVILKQYHLVHSEITTC---HDVVD 265
 DB 129 VHPVRQRQRTVVIARTCTVLIAWAFAPLALSLPTALSRDVTINHPNTVCGTLDKHEL-- 186
 QY 266 ACESPSRFYFYFVSL-----AFGFLIPFVIIIFCYTTLIHLKSKDRIWLG-----YI 315
 DB 187 -----SHVLVAIGLMKSVLGFLLIPFVIIIVTCYCLIGALLAARRVQSSRSRGDEV 237
 QY 316 KAVLLILVITICFAPTNILVIHANYHYNTDSLYFWYLIAL----- 359
 DB 238 QMLAAVVLAFPLCWPHQI-----PFHFMVLALLKVIENCPPLDIIIDTA 281
 QY 360 -----CLGSNSCLDPLFY 373
 DB 282 LPFTICIAFNNSCNPILY 300
 RESULT 18
 Q9ERK9 ID Q9ERK9 PRELIMINARY; PRT; 328 AA.
 AC Q9ERK9
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE P2Y6 receptor (Hypothetical 36.7 kDa protein).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=129/SVEV;
 RC MEDLINE=21160052; PubMed=11259526;
 RX Lazarowski E.R., Rochelle L.G., O'Neal W.K., Ribeiro C.M.P.,
 RA Grubb B.R., Zhang V., Harden T.K., Boucher R.C.;
 RT "Cloning and functional characterization of two murine uridine
 RT nucleotide receptors reveal a potential target for correcting ion
 RT transport deficiency in cystic fibrosis gallbladder."
 RL J. Pharmacol. Exp. Ther. 297:43-49(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF298899; AAG24619.1; -;
 DR EMBL; BC027331; AAH27331.1; -;
 DR HSSP; P34996; 1DDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPS.
 DR PROSITE; PS00262; G-PROTEIN RECEPT_F1_2; 1.
 KW Hypothetical protein; Receptor.
 SQ SEQUENCE 328 AA; 36721 MW; 00F9DF5ADADF903E CRC64;
 Query Match 15.3%; Score 326; DB 11; Length 328;
 Best Local Similarity 27.6%; Pred. No. 6.5e-20;
 Matches 86; Conservative 55; Mismatches 109; Indels 62; Gaps 12;
 QY 104 RSLSTQVIPAIIYILLFVVGVPSPNIVTLWKL--SLRTKTSISLIVFHTNLAIADLFLFCVTL 161
 DB 21 REDFKRLLLTPYSVYVGLPLNCVIAQICASRRLTRLS-AVYTLNLALADLWYACSL 79
 QY 162 PFKIAVHLNGNNWVGEVNCRTTTFVFGNMYCAILLITCMGINRYLATAPFP-TVOKLP 220
 DB 80 PLLIINYARGDHWPGDGLACRFVRFYANLHGSILFLTCISFQRYLGICHPLASWHKRG 139
 QY 221 KRSPSLMCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVDVACESPS--SPRFY-Y 277
 DB 140 GREAAMVGVVLAATAVTAQCLPTAVFAATG--IQNRITVCYDL-----SPPISTLYLPY 192
 QY 278 FVSLAFFGLIPFVIIIFCYTTLIHLKSKDRIWLG-----YIKAVLLILVIF 325
 DB 193 GMLATVIGFLPFIALLACYCRMARRLCQD----GPAGPVAQERRSKAARMVAVVAVF 248
 QY 326 TICCAPTNILVIHANYHYNTDSLYF-----MYLIALCLGSLNSC 367
 DB 249 AILSFLP-----FHITKAYLAVRSTPGVSCPVLETPFAAAYKTRPPASVNSV 295
 QY 368 LDPFL-YFVMSK 378
 DB 296 LDPILFYFTQOK 307
 RESULT 19
 Q8QFR5 ID Q8QFR5 PRELIMINARY; PRT; 353 AA.
 AC Q8QFR5
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE CXCR2 like protein.
 GN CXCR2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]

QY 349 -----DSLFYMLIALCLGSLNSCLDPFLY-FVMSKVVDQL 383
DB 259 RGGKVIETCELRSIDVALVYVQAMAFACAINPILYAFIGKFRNQL 306

RESULT 22

O46685 PRELIMINARY; PRT; 361 AA.
AC O46685; (T-EMBLrel. 06, Created)
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Orphan G protein-coupled receptor BRGRIB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Ferrer-Martinez A., Felipe A., Casado J., Pastor-Anglada M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; U88367; AAC05612.1; -;
DR EMBL; U88366; AAC05611.1; -;
DR HSPSP; P02699; IPR8.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR01564; OGR1RECEPTOR.
DR PRINTS; PR01649; PSYCHOSINER.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SEQUENCE 361 AA; 40905 MW; 4802E77C9A45DEE1 CRC64;

Query Match 15.0%; Score 319.5; DB 6; Length 361;
Best Local Similarity 26.7%; Pred. No. 2.6e-19;
Matches 76; Conservative 65; Mismatches 107; Indels 37; Gaps 8;
QY 113 PAIYILFVVGPSNIVTWKLSIRTKSIS-LVIFHTNLAIDLFCVTLPFKIAHYLNG 171
DB 25 PVVYVVLVVGFPANCLSLYGYLQIKARNELGVLCNLTVADLFYICSLPFWLQVLOH 84
QY 172 NNWVGEVMCRITTVFVYGNMYCAILITCMGINRYLATAHPTFYOKLPKRSFSLMCGI 231
DB 85 DHWSHDDLSCQVCGILLIYENIYISVGFLLCCISIDRYLAVAHAPRFHQFRTLKAAMGVSA 144
QY 232 VVMVFLYMLPFVILKQEHVHSEITTCDDVDAC-----ESPSSFRFYFVSLAFF 284
DB 145 IWKELTSTI-----YFLMHEEVVEDADHRVCFEHPYLPBPRGINYRFL----V 192
QY 285 GFLIPFVIIFCYTTLIHKLK-----SKDRIMLGKIAKAVLLIIVIFTCFAPTNIL 336
DB 193 GFLFPICLLASRGILRAVRSHGCTQSKRQIQ-----RLVLTSTVIFLACPLPHVLL 248
QY 337 VIHANYYYHNTD---SLYPMYIALCLGSLNSCLDPFLYFVMSK 378
DB 249 LVR--SLWESSCDFAKGIFNAYHFSLLTSTFNCAVDADPVLVCFVSE 291

RESULT 23

O9UNW8 PRELIMINARY; PRT; 380 AA.
AC O9UNW8;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE G protein-coupled receptor.
GN G2A.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Weng Z., Witte O.N.;
RT "A DNA damage and stress inducible G protein-coupled receptor blocks
RT cells in G2/M";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083955; AAD47380.1; -;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR01563; G2ARECEPTOR.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; UNKNOWN 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 380 AA; 42499 MW; 6DE63D17275ECD74 CRC64;

Query Match 15.0%; Score 319.5; DB 4; Length 380;
Best Local Similarity 25.5%; Pred. No. 2.7e-19;
Matches 82; Conservative 67; Mismatches 131; Indels 41; Gaps 10;
QY 73 GWTGATTIKAECPEDS--ISTLVNNTATIGYLRSSLSTQVTPAIYIILLFVVGVSNIYT 130
DB 9 GYNGNATPVTTTAPWASLGLSAKTCNNVSF-----EESRVLVVVYSVACTILGVPA 63
QY 131 LMKLSIRT-KSISLVIFHTNLAIDLFCVTLPFKIAHYLNNWVFGVMCRITTVPY 189
DB 64 AMLALLQVGNLVALYLLCLALCELLYGTGLPLVIYIRNQHRWTLGLLACKVTAYIFP 123
QY 190 GNMVCAILITCMGINRYLATAHPTFYOKLPKRSFSLMCGIWMVFLYMLPFVILKOE 249
DB 124 CNIYVSIILFLCCISCDRFVAVVVALESRRRRRTAILISACIFILVGVHVP--VFQTE 181
QY 250 YHLVHSEITTCDDVDACESPSSFRFYFVSLAFFGLIPFVIIFCYTTLIHKLKSKDR 309
DB 182 -----DKETCFDML--QMSRSIAGYVYARFT-VGFAIPLSIIAFTNHRIFRSIKQS-- 229
QY 310 IWLGVYKA-----VLLILVIFTCFAPTNILVIHANYYYHNTD-----SLY 352
DB 230 --MGLSAAQAKAKVHSAIAVVVIFLVCFAPYHLLVLLVKAASFYRGDRNMCGLLEELY 287
QY 353 FMYIALCLGSLNSCLDPFLY 373
DB 288 TASVFLCLSTVNGVADPIY 308

RESULT 24

O9TTY5 PRELIMINARY; PRT; 342 AA.
AC O9TTY5;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2001 (T-EMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Platelet-activating factor receptor.
GN PAFR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang W.S., Diehl J.R., Roudebush W.E.;
RT "Partial Sequence of Bovine Platelet-Activating Factor Receptor
RT Gene";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA He B., Tiemann U., Kanitz W., Weikard R., Laurent P., Schwerin M.,
RA Schmidt P.;
RT "Molecular characterization of bovine platelet-activating factor

RT receptor transcripts and their detection in different tissues of
 RT cattle.",
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF187321; AAF01439.2; -
 DR EMBL; AJ295321; CAC43290.1; -
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR01559; DUFFYANTIGEN.
 DR PRINTS; PR0237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G PROTEIN RECEPTOR F1.1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 342 AA; 39691 MW; 7C4236205AE937C9 CRC64;

Query Match 14.9%; Score 318; DB 6; Length 342;
 Best Local Similarity 27.5%; Pred. No. 3.2e-19;
 Matches 82; Conservative 50; Mismatches 120; Indels 46; Gaps 8;

QY 111 VIPAIYILLFVVGVPNSIVTLW---KLSLRKTSISLVIFHTNLAIADLLFCVTLPPKIA 167
 Db 17 LPPIFYISIVFVLGVANSYLVWVFARLYPSKKEIKIFMWNLTWADLLFLVTLPLWIVY 76
 QY 168 HLNNGNNVFEVGMCRITTVVFGNMYCAIILITCMGINRYLATAHPFTYQKLPKRSFSL 227
 Db 77 YNQGDIWLPKFLCNLAGCFFINITYCSVAFLAVITYNRFOAVTRPIKTAQATTKRGFL 136
 QY 228 MCGIVVMVFLYMLPFVIL---KQYHLVHSEITTCDDVDVADACESPSSFRFYFVSLAF 283
 Db 137 LSLIIWVSIVGAASYFFVLDSTNRPKNGTSANITRCFEHYEKGSIPLVLIHFLV---- 192
 QY 284 FGLPLFPVILFCYTLIIHLKSKD-----RIWLGVIKAVLLILVIFTICFAP 331
 Db 193 FSFFLVFLIILFCNLVIIRTLTQQVQIQORNAEVRKRLWM-----VCTVLAFFVFCVP 247
 QY 332 TMIILVIH-----ANYVYHNTD---SLVFWYLIACLSLNSCLDPPFLYFVMSK 378
 Db 248 -----HHLVQLPWTLAELGFQDTHQAINDAHQVTLCLLSTNCVLDPIIYCFLTK 298

RESULT 25

Q9GK76 PRELIMINARY; PRT; 342 AA.
 AC Q9GK76;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Platelet-activating factor receptor.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang W., Diehl J.R., Piumi F.;
 RT "Partial Sequence of Caprine Platelet-Activating Factor Receptor Gene."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF302764; AAG39882.1; -
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR01559; DUFFYANTIGEN.
 DR PRINTS; PR0237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G PROTEIN RECEPTOR F1.1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 342 AA; 39612 MW; 39D9379AED60F04E CRC64;

Query Match 14.8%; Score 317; DB 6; Length 342;
 Best Local Similarity 27.8%; Pred. No. 4e-19;
 Matches 84; Conservative 46; Mismatches 118; Indels 54; Gaps 9;

QY 111 VIPAIYILLFVVGVPNSIVTLW---KLSLRKTSISLVIFHTNLAIADLLFCVTLPPKIA 167

Db 17 LPPIFYISIVFVLGVANSYLVWVFARLYPSKKEIKIFMWNLTWADLLFLVTLPLWIVY 76
 QY 168 HLNNGNNVFEVGMCRITTVVFGNMYCAIILITCMGINRYLATAHPFTYQKLPKRSFSL 227
 Db 77 YNQGDIWLPKFLCNLAGCFFINITYCSVAFLAVITYNRFOAVTRPIKTAQATTKRGFL 136
 QY 228 MCGIVVMVFLYMLPFVILKQYHLVHSE-----ITTCDDVDVADACESPSSFRFYFV 279
 Db 137 LSLIIWVSIVGAASYFFVLDSTNRPKNGTSANITRCFEHYEKGSIPLVLIHFLV 192
 QY 280 SLAFGLFPLFPVILFCYTLIIHLKSKD-----RIWLGVIKAVLLILVIFTI 327
 Db 193 -----FSFFLVFLIILFCNLVIIRTLTQQVQIQORNAEVRKRLWM-----VCTVLAFFV 243
 QY 328 CFAPNTMIILVIH-----ANYVYHNTD---SLVFWYLIACLSLNSCLDPPFLYFV 376
 Db 244 CFPV-----HHLVQLPWTLAELGFQDTHQAINDAHQVTLCLLSTNCVLDPIIYCFL 296
 QY 377 SK 378
 Db 297 TK 298

RESULT 26

Q9NOU1 PRELIMINARY; PRT; 359 AA.
 AC Q9NOU1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Angiotensin II type 1 receptor.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ADRENAL CORTEX;
 RA Bird I.M., Millican D.S.;
 RT "Isolation of a Full Length Ovine Angiotensin II Type-1 Receptor (AT1-R) cDNA."
 RL Endocr. Res. 0:0-0(2000).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AF254119; AAF66063.1; -
 DR HSP; P34996; IDDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G PROTEIN RECEPTOR F1.1; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
 SQ SEQUENCE 359 AA; 41062 MW; C07010EDB81110EB CRC64;

Query Match 14.7%; Score 315; DB 6; Length 359;
 Best Local Similarity 27.6%; Pred. No. 6.2e-19;
 Matches 85; Conservative 50; Mismatches 103; Indels 70; Gaps 8;

QY 111 VIPAIYILLFVVGVPNSIVTLW---KLSLRKTSISLVIFHTNLAIADLLFCVTLPPKIA 166
 Db 30 MIPTLYSIIFVVGFGSLWIVIVFYMKLKTVA---SVFLNLALADLCFLTLPLWAV 86
 QY 167 YHLNGNNVFEVGMCRITTVVFGNMYCAIILITCMGINRYLATAHPFTYQKLPKRSFSL 226
 Db 87 YTAMEYRWPFGNYLCKIASGSVSNLYASVFLTLCLSDRYLAIVHPMKSRRLRRTMLVAK 146
 QY 227 LMCGIVVMVFLYMLPFVILKQYHLVHSEITTCDDVDVADACESPSSFRFY----- 277
 Db 147 VTCIIWLAGLASPTIIRNVFFIENTNITVC-----APHYSQNSTLPV 193
 QY 278 --FVSLAFFGLIPFVIIIFCYTLIIHLKLS-----KDIRWLGVIKAVLLILVIF 325

Db 194 GLGLTRNKLGLFPFLPILTLTYLWTKLKAYEIQNKPRKDDIF---KIILAIVLFF 249
Qy 326 TICFAPTNILVHANYHYHNTDSLYFMYL-----IALCLGSLNSCLDP 370
Db 250 FFSWPHQIFTFN-----DVLQIQLRDCKIEDIVDTAMPITICLAYFNCLMP 299
Qy 371 FLYFVMSK 378
Db 300 LFYGFGLK 307

RESULT 27
Q8QG04 PRELIMINARY; PRT; 390 AA.
AC Q8QG04; Query Match 14.6%; Score 311; DB 13; Length 390;
DT 01-JUN-2002 (TrEMBLrel. 21, Created) Best Local Similarity 28.4%; Pred. No. 1.5e-18;
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update) Matches 80; Conservative 64; Mismatches 116; Indels 22; Gaps 7;
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update) Type five-like somatostatin receptor.
DE Type five-like somatostatin receptor.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Oetariophysi; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Peter R.E.;
RT "Cloning and expression of a type five-like somatostatin receptor in goldfish."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF252879; AAM18805.1; --
KW Receptor.
SQ SEQUENCE 390 AA; 44136 MW; EC12119A4B6CF9A8 CRC64;

Query Match 14.6%; Score 311; DB 13; Length 390;
Best Local Similarity 28.4%; Pred. No. 1.5e-18;
Matches 80; Conservative 64; Mismatches 116; Indels 22; Gaps 7;
Qy 108 STQVIPAIIYLLFVGVPSNIVTLWKLRLTK-SISLVIFHTNLATADLLFCVTLPPFKIA 166
Db 39 STKILAVILWVFGVGLGSLAIFVFLRYTKMTATNNMYILNLAVADELYLGLPFLTA 98
Qy 167 YHLGNWVFGVGMCRITTVFVGNMYCAIILTCMGINRYLATAHPTFYQKLPKRSFSL 226
Db 99 HNMVG-YWPGNFCILMKWTDISQSTFTCLTVMSIDRYMAVHPISARWRPRVAK 157
Qy 227 LMGIVVMVFLYMLPFVILKQYHLVHSEITTCCHDVVDACSPSSFRFYFVSLAFFGF 286
Db 158 VINSWVWALSCLTLFVII---YCDVQPELNTCN--LSWPEPRDVMSTAFILYTMALGF 211
Qy 287 LIPFVILIFCYTTLIHLKLS-----KDRWLGYIKAVLLILVIFTICFAPTNIILV 337
Db 212 FPELMVICLLYLIVIKVSASARAGLSKRGSEKVKTRMVIIVVFLVCLPFFIINI 271
Qy 338 IHANVYHNT--DSLVMYLIACLGSLNSCLDPFLYFVMS 377
Db 272 LNLISLPNSLMTGIFLTVI---LTYVNSCANPLLYSFLS 310

RESULT 28
Q9XSD4 PRELIMINARY; PRT; 342 AA.
AC Q9XSD4; Query Match 14.5%; Score 310; DB 6; Length 359;
DT 01-NOV-1999 (TrEMBLrel. 12, Created) Best Local Similarity 27.5%; Pred. No. 1.6e-18;
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update) Matches 81; Conservative 50; Mismatches 124; Indels 40; Gaps 8;
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update) Type platelet-activating factor receptor.
DE Platelet-activating factor receptor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]

RP SEQUENCE FROM N.A.
RT Diehl J.R., Murphy K.E., Roudabush W.E.;
RL "Porcine (Sus scrofa) platelet-activating factor receptor DNA";
DR Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF124054; AAD28739.2; --
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1; UNKNOWN 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW Receptor.
SQ SEQUENCE 342 AA; 39435 MW; 830EA7542B528DAC CRC64;

Query Match 14.5%; Score 310; DB 6; Length 342;
Best Local Similarity 27.5%; Pred. No. 1.6e-18;
Matches 81; Conservative 50; Mismatches 124; Indels 40; Gaps 8;
Qy 111 VIPAIYILLFVGVPSNIVTLW---KLSLRKTSISLVIFHTNLAIADLLFCVTLPPFKIAY 167
Db 17 LFPFIYSIIFVLGVIANSVLWVPARVYPSKLNKIKIFMLNLTMADLLFLVTLPLWIY 76
Qy 168 HUNGNNWVFGVGMCRITTVFVGNMYCAIILTCMGINRYLATAHPTFYQKLPKRSFSL 227
Db 77 YYHEGNWILPKFLCNLAGCFEINTYCSVFLAVIYNRFQAVTRPIKTAQATTKRGIS 136
Qy 228 MCGIVVMVFLYMLPFVILKQYHLVHS---EITTCCHDVVDACSPSSFRFYFVSLAF 283
Db 137 LSLIIVAMVAAASYFFVLDSTNIELSKTGAGNLTRCFEHYKSGMPVLIHIFLV--- 192
Qy 284 FGLIPFVILIFCYTTLIHLKLSKD-----RWLGYIKAVLLILVIFTICFAP 331
Db 193 FSEFLVFLVILFVLCNLVIITLTLTQSVQMRNAEVKRRALWM---VCTVLAVFICFVP 247
Qy 332 TNII-----LVTHANYHYHNTDSLYFMYLIACLGSLNSCLDPFLYFVMSK 378
Db 248 HHIVQLPWTABLPGSQGNFHOAINDA---HQITCLLSTNCVLDPIIYCFITK 298

RESULT 29
Q9GLN9 PRELIMINARY; PRT; 359 AA.
AC Q9GLN9; Query Match 14.5%; Score 310; DB 6; Length 359;
DT 01-MAR-2001 (TrEMBLrel. 16, Created) Best Local Similarity 27.5%; Pred. No. 1.6e-18;
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update) Matches 81; Conservative 50; Mismatches 124; Indels 40; Gaps 8;
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update) Type Angiotensin II type-1 receptor.
DE Angiotensin II type-1 receptor.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20469400; PubMed=11013071;
RA Dufour C., Casane D., Denton D., Wickings J., Corvol P.,
RA Jeunemaitre X.;
RT "Human-Chimpanzee DNA sequence variation in the four major genes of the renin-angiotensin system."
RL Genomics 69:14-26(2000).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF193445; AAG28410.1; --
DR HSP; P34996; 1DDO.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 359 AA; 41033 MW; 266133A7A3E911A6 CRC64;

Query Match 14.5%; Score 310; DB 6; Length 359;

```
Best Local Similarity 26.4%; Pred. No. 1.6e-18;
Matches 85; Conservative 54; Mismatches 131; Indels 52; Gaps 8;

QY 87 EDSISTLVHNNATIGYLRSSLSQVIPAIIYLLFVVGVPSPN---IVTLWKLRLRTKSTIS 142
Db 8 EDGIRKIQDDCPKAG--RHNYIFVMIPTLYSIIFVVGISLVVIVFYMKLKTVA-- 63

QY 143 LVIFHTNLAIADLLFCVTLPPFKIAHYHLNGNNWVFGVMCRITTVVFGNMYCAIILTCM 202
Db 64 -SVFLNLALADLCFLTLPLWAVYTAMEYRPPFGNYLCKIASASVSFNLYASVFLLTCL 122

QY 203 GINRYLATAHPPTYQKLPRKSFSLMCGIWMVFLYMLPFVILKQEHYLVHSEITCHD 262
Db 123 SIDRYLAIVHPMKSLRRRTMLVAKVTCIIIIWLLAGLASLPAIIHRNVFFIENTNITVC-- 180

QY 263 VVDACESPSSFRFY-----FVSLAFPGFLIPFVFIIFCVTTLIHLKSKDRW 311
Db 181 -----AFHYESQNSTLPIGLTKNLTGFLPFLIILTSYTLWKLKAYEIQ 229

QY 312 LG-----YKAVLLILVITFCFAPTNIIL-----VIHANYYYHNTDSLYFMYL 356
Db 230 KKNPRNDIFKIIAIVLFFFSWIPHOIFTFDLVLIQLGIIRDCRI----ADIVDTAMP 285

QY 357 IALCLGSLNSCLDPPFLYFVMSK 378
Db 286 ITICIAFYNNCLNPLFYGLGK 307

RESULT 30
Q8V3F0 PRELIMINARY; PRT; 370 AA.
AC Q8V3F0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE SPV146 G protein-coupled CC chemokine receptor-like protein.
GN SPV146.
OS Swinepox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Suipoxvirus.
OC NCBI_TaxID=10276;
RN [1]_TaxID=10276;
RP SEQUENCE FROM N.A.
RC STRAIN=17077-99;
RX MEDLINE=21624277; PubMed=11752168;
RA Afonso C.L., Tulman E.R., Lu Z., Zaak L., Osorio F.A., Balinsky C.,
RA Kutish G.F., Rock D.L.;
RT "The genome of swinepox virus.";
RL J. Virol. 76:783-790(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=17077-99;
RA Afonso C.L., Tulman E.R., Lu Z., Balinsky C., Osorio F.A., Zaak L.,
RA Kutish G.F., Rock D.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF410153; AAL69885.1;
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHOPOSP.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 370 AA; 43164 MW; 655AF5DDDD065D1D1 CRC64;

Query Match 14.5%; Score 309.5; DB 12; Length 370;
Best Local Similarity 27.2%; Pred. No. 1.9e-18;
Matches 87; Conservative 68; Mismatches 130; Indels 35; Gaps 12;

QY 80 TKAECPEDSISTLVHNNATIGYLRSSLSQVIPAIIYLLFVVGVPSPNIVTLWKLRLRT 138
Db 34 TINNDITSSSVIKAFNNCT--FLEDTKHYIIV--IHLIULGSGTGNIFVVSIIAFKRN 89

QY 139 KSISLVIFHTNLAIADLLFCVTLPPFKIAHYHLNGNNWVFGVMCRITTVVFGNMYCAIIL 198
```

```
Db 90 KSIT-DIVILNLSMDCIFVFOIPFIVYSKL--DQWIFGNILCKIMSVLYVYVGFPSNMPF 146
QY 199 LTCMGINRYLATAHPPTYQKLPRKSFSLMCGIWMVFLYMLPFVILKQEHYLVHSEIT 258
Db 147 ITLMSIDRYFAIVHPDIKQPYRTKIGILMCCSAMLSSILSSPVSKLYENIPHSKDIY 206

QY 259 TCHDVVDACESPSSFRFYFVSLAFPGFLIPFVFIIFCY-----TTLI-----HKLKSKD 308
Db 207 QC-TLTNENDSIATIKRLMQIEITILGFLIPIIFVVCYRIFFTVVRNRKRYKS-- 263

QY 309 RIWLGYIKAVLLILVITFCFAPTNIILVI-----HHANYYYHNTDSLYFMYLIAL--C 360
Db 264 -----IKVLMIVVCSLICWIPLYVLMATIVSLYTSNIFRHLCLYLNLAYAITFSET 317

QY 361 LGSLSNCLDPPFLYFVMSKV 380
Db 318 ISLARCCINPIIYTLIGEHV 337

RESULT 31
Q8TBK4 PRELIMINARY; PRT; 359 AA.
AC Q8TBK4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Angiotensin receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022447; AAH22447.1;
KW Receptor.
SQ SEQUENCE 359 AA; 41087 MW; AF90BE7F21E911BA CRC64;

Query Match 14.4%; Score 308; DB 4; Length 359;
Best Local Similarity 26.7%; Pred. No. 2.4e-18;
Matches 87; Conservative 56; Mismatches 123; Indels 60; Gaps 10;

QY 87 EDSISTLVHNNATIGYLRSSLSQVIPAIIYLLFVVGVPSPN---IVTLWKLRLRTKSTIS 142
Db 8 EDGIRKIQDDCPKAG--RHNYIFVMIPTLYSIIFVVGISLVVIVFYMKLKTVA-- 63

QY 143 LVIFHTNLAIADLLFCVTLPPFKIAHYHLNGNNWVFGVMCRITTVVFGNMYCAIILTCM 202
Db 64 -SVFLNLALADLCFLTLPLWAVYTAMEYRPPFGNYLCKIASASVSFNLYASVFLLTCL 122

QY 203 GINRYLATAHPPTYQKLPRKSFSLMCGIWMVFLYMLPFVILKQEHYLVHSEITCHD 262
Db 123 SIDRYLAIVHPMKSLRRRTMLVAKVTCIIIIWLLAGLASLPAIIHRNVFFIENTNITVC-- 180

QY 263 VVDACESPSSFRFY-----FVSLAFPGFLIPFVFIIFCVTTLI-----IH 302
Db 181 -----AFHYESQNSTLPIGLTKNLTGFLPFLIILTSYTLWKLKAYEIQ 229

QY 303 KKLKSK-DRIWLGVIKAVLLILVITFCFAPTNIIL-----VIHANYYYHNTDSLY 352
Db 230 KKNPRNDIF-----KIIMAVLFFFSWIPHOIFTFDLVLIQLGIIRDCRI----ADIVD 281

QY 353 FMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 282 TAMPITICIAFYNNCLNPLFYGLGK 307

RESULT 32
Q9UQ06 PRELIMINARY; PRT; 369 AA.
ID Q9UQ06
```


Matches	77;	Conservative	57;	Mismatches	118;	Indels	37;	Gaps	9
Qy	112	IPAIYILLFVVGVP	SNIIVTLKWSLRTKSISLV-IFHTNLIAIADLLFCVTLPFKIAYHLN	170					
Db	41	VPIAYSLSICVFGLG	NIMVMVMTAFYKKARSMDTVLLNMAITDILFVLTLPPWAVTHAT	100					
Qy	171	GNWVFGVWCRTITT	VVFGNMVCAIIILTCMGINRYLATAHPTTYOKLPRSF--SLLM	228					
Db	101	-NTWFSDALCKLMKG	TAYVNFNCMLLACISMRIYAIVQATKSPRVSRTLTHSKVI	159					
Qy	229	CGIIVMWVFLYMLP	FVLIKOEYLHVHSEITTHCDVVDAESPSSFRFYFVSLAFCFLI	288					
Db	160	CVAWFIISIISSTP	TFINKYQLQRDV--CEPRYSVSEPITKULLGLGLELFFGGFT	217					
Qy	289	PFIUIIFYC----	TTLIHKLSKDRIWLGYIKAVLLIIVTICTFAPTNIILVIHHAN--	342					
Db	218	PLLFMFVCYLI	FIKTIVQAQNSKH---RAIRVVIAVLVFLACOIPHNVLLVTAVNG	274					
Qy	343	-----YYYINTDSLY	PMYLIALCLGSINSLCDPPLY-FVMSK	378					
Db	275	KVARSCSTEKVLA	YTRN-----VAEVLAFHLHSLNDPVLYAFIQG	314					
 RESULT 36									
Q9MYJ8									
ID	Q9MYJ8	PRELIMINARY;	PRT:	355	AA.				
AC	Q9MYJ8								
DT	01-OCT-2000	(TEMBLrel. 15, Created)							
DT	01-OCT-2000	(TEMBLrel. 15, Last sequence update)							
DT	01-JUN-2001	(TEMBLrel. 17, Last annotation update)							
DE	Chemokine receptor.								
GN	CCRL1.								
OS	Callithrix jacchus (Common marmoset).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.								
OX	NCBI_TaxID=9483;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=20153429; PubMed=10686294;								
RA	Liang M., Rosser M.P., NG H.P., May K., Bauman J.G., Islam I.,								
RA	Ghanam A., Kretschmer P.J., Pu H., Dunning L., Snider R.M.,								
RA	Morrissey M.M., Heeslgesseer J., Perez H.D., Horuk R.;								
RT	"Species selectivity of a small molecule antagonist for the CCR1								
RT	chemokine receptor.";								
RL	Eur. J. Pharmacol. 389:41-49(2000) .								
DR	EMBL; AF127528; AAF36453.1; -								
DR	InterPro; IPRO00276; GPCR_Rhodpsn.								
DR	Pfam; PF00001; 7tm_1; 1.								
DR	PRINTS; PR00237; GPCRHDOPSN.								
DR	PROSITE; PS00237; G PROTEIN RECEPT_F1_1; UNKNOWN_1.								
DR	PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.								
KW	Receptor.								
SQ	SEQUENCE	355	AA;	40928	MW;	2B01C47E9874A2C1	CRC64;		
 Query Match									
Best Local Similarity 14.3%; Score 305; DB 6; Length 355;									
Matches 86; Conservative 67; Mismatches 124; Indels 44; Gaps 12;									
Qy	75	TGATTTIKACPDSISTLHN	NATIGYURS--SLSTQVIPAIYILLFVVGVP	SNIIVTLW	132				
Db	3	TSATTE-----DYDMITEFD	YSDATPKHANERAFAGKLLPPLYSLVFVIGLVGNILVVV	57					
Qy	133	KL--SLRTKSISLVI	FHTNLIAIADLLFCVTLPFKIAYHLNNVWVGVWCRTITVVYFG	190					
Db	58	VLVOYKRLKNMT-SI	YLLNLIAISDLFLFLFPWISYOLK-TDWVFNAMCKVLSGFYYT	115					
Qy	191	NMYCAIIILT	CMGINRYLATAHPTTYOKLPRKSFSLLMCGIVWMVFLYMLP-FVILKQE	249					
Db	116	GLYSEIFFILLTID	RYLAIVHAVFAIRARTVTFVGITSIIIVLAISLPGLYFAKTO	175					
Qy	250	YHLVHSEITTC----	HDDVDAESPSSFRFYFVSLAFFGLIPFVIIICYTTLI---	301					
Db	176	WEITH---RTCSLHFPH	-----ESRQEWKLFQALKNLGLVPLVLMVICVYTGIIKIL	226					

Query Match	14.3%	Score 305.5;	DB 11;	Length 367;
Best Local Similarity	26.6%;	Pred. No. 4.1e-18;		

```

QY 302 ----HKLKSDRIWLGVIKAVLLILVITFCAPTNIILVI-----HHANYYYHNTDSLY 352
DB 227 LRRPNKKK-----AVRLIFVIMIFLFWTPYNNLTLLISVFQDFLFTYCEQGRQLD 280

QY 353 FMYLIALCLGSLNSCLDPFLY 373
DB 281 LAIQVTEMIAYTHCCVNPVY 301

RESULT 37
Q9BDS6 PRELIMINARY; ~ PRT; 343 AA.
AC Q9BDS6;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Orphan seven transmembrane receptor STRL33.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21134756; PubMed=11242524;
RA Wade-Evans A.M., Russell J., Jenkins A., Javan C.;
RT "Cloning and sequencing of cynomolgus macaque ccr3, gpr15, and strl33:
RT potential coreceptors for HIV type 1, HIV type 2, and SIV."
RL AIDS Res. Hum. Retroviruses 17:371-375(2001).
DR EMBL; AF291671; AAK25742.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; LYMPHOTACTNR.
DR PROSITE; PS00237; G PROTEIN RECEPTOR_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G PROTEIN RECEPTOR_F1_2; 1.
KW Receptor; Transmembrane.
SQ SEQUENCE 343 AA; 39472 MW; 0961328F948E7784 CRC64;

Query Match 14.3%; Score 304.5; DB 6; Length 343;
Best Local Similarity 26.9%; Pred. No. 4.6e-18;
Matches 87; Conservative 62; Mismatches 108; Indels 67; Gaps 14;

QY 112 IPAIYLLFVVGVPNSIVTLWLKSLRTKSLV-IFHTNLAIADLLFCVTLDPFKIAYHLN 170
DB 35 LPCMLVFLVGLVNSLVLSIFVHKQLSLTDVFLNPLADLVFVCTLPFWT--YAG 92

QY 171 GNNVFGVNCRTTIVVFGNMYCAIILTCMGINRYLATAHPF-TYQKLPKR-SFSLM 228
DB 93 IHEWIFGVCKTLLGVYTNFYTSMLITCTIVDRFVVVKATKAYNOQAKMTGWKVI 152

QY 229 CGIWWNVFLYMLPFVILKQYHL-----VH-SEITTHDVVDACESPSSFRFYFVSL 281
DB 153 CLLIWISSLVSLPQIYVGNVFNLDKLCYHDEEIST--VVLATQMT----- 198

QY 282 APTGELIPFVIIIFCYTLLHLK-----SKDRIWLGVIKAVLLILVITFCAPTNIIL 336
DB 199 --LGFPLPMLAMIVCVSVIKTLHLHAGGFKHR-----SLKIIFLVMAVLLTQTPNVLK 252

QY 337 VIHANYYYHNTDSLVMYLIACLGSLNSCLDPFLY-FWMSKVVDQLNPXSAMARPLXR 395
DB 253 LIRSTRWEYAMTSFHYTIIVTAIYLRACLPVLYAFVSLK----- 296

QY 396 PRDIW---BDI-----HAW 407
DB 297 -RKNFWKLVRDGLCLPYLGVSHOW 319

RESULT 38
O42324 PRELIMINARY; PRT; 383 AA.
ID O42324
AC O42324;
```

```

DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Mu-opioid receptor.
OS Catostomus commersoni (White sucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Catostomidae; Catostomus.
OX NCBI_TaxID=7971;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=CNS;
RC MEDLINE=97368346; PubMed=9223341;
RA Darlison M.G., Greten F.R., Harvey R.J., Kreienkamp H.J., Stuenkel T.,
RA Zwiers H., Lederis K., Richter D.;
RT "Opioid receptors from a lower vertebrate (Catostomus commersoni):
RT rectifying potassium channel (GIRK1), and evolution-gated inward-
RT sequence, pharmacology, coupling to a G-protein-gated inward-
RL Proc. Natl. Acad. Sci. U.S.A. 94:8214-8219(1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; Y10904; CA871843.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G PROTEIN RECEPTOR_F1_1; 1.
DR PROSITE; PS00262; G PROTEIN RECEPTOR_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 383 AA; 43232 MW; F58838B57A107305 CRC64;

Query Match 14.2%; Score 304; DB 13; Length 383;
Best Local Similarity 28.9%; Pred. No. 5.7e-18;
Matches 94; Conservative 61; Mismatches 108; Indels 62; Gaps 14;

QY 88 DSISTLHVNNTATIGVLRSSLSLSTQ-----VIPATVILLFVVGVPNSIVTLWLKSLRT 138
DB 31 NSTSFLNN-----GSSRSDSTDEQDKTPVIAIIITLTSIVCVGLGNVLMVYIIRYT 86

QY 139 K--SISLIVPHTNLAIADLLFCVTLDPFKIAYHLNKNVFGVNCRTTIVVFGNMYCA 195
DB 87 KMKTAIYIP--NLALADALATSLTPQSVNYLMG-TWPGDVGVCKIWMSDIYNMFTS 143

QY 196 ILILTCMGINRYLATAHPFTYQ--KLPRKSFSLMCGIWMVVMVFLYMLPFVILKQYHLV 253
DB 144 IFTLTMSIDRYIAVCHPVKALDPRTPNAKIVNVCN--WILSSAIGLPVMVMAS----- 196

QY 254 HSEITTHDVVDACESP--SSFRF-----YFVSLA-----FFGFLIPFVILIICY 297
DB 197 ----TT-----IENQNSPLQVSNFDCITLLFPHPPWYMETLLKTCVFILAFIMPVLIITCY 248

QY 298 TTLHLKLS-----KDRIWLGVIKAVLLILVITFCAPTNIILVIHANYYYHNT 348
DB 249 GLMIURLKSVMLSSGSKEDRNLRITRWLVNVAVFIICTWPIHIFVII-KALVTIPNS 307

QY 349 DLSLEYMYLIACLGSLNSCLDPFLY 373
DB 308 LFQVTWHEFIALGYTNSCLNPVLY 332

RESULT 39
Q9JLZ0 PRELIMINARY; PRT; 362 AA.
AC Q9JLZ0;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Orphan chemokine receptor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC TISSUE=MUSCLE;
RA Xie P., Fu A.K.Y., Ip N.Y.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AF118816; AAF34338.1; -.
DR HSSP: P34996; 1DDO.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm1.1.
DR PRINTS: PR00237; GPCRHHODOPGN.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 362 AA; 41650 MW; C47D1DD678697CFD CRC64;

Query Match      14.2%; Score 303.5; DB 11; Length 362;
Best Local Similarity 25.8%; Pred. No. 5.9e-18;
Matches 73; Conservative 62; Mismatches 117; Indels 31; Gaps 8;

QY 115 YILLFVVGVPNIVTLW-KLSLRKTSISLVPHNTLAIDLFCVTLFPKIAVHLNGNN 173
Db 115 YILLFVVGVPNIVTLW-KLSLRKTSISLVPHNTLAIDLFCVTLFPKIAVHLNGNN 173
QY 50 IYIFIEFVIGMIANSVVVWNIQAKTTGYDTHCYILNLAIADLWVITIPVWVSVLQHNQ 109
Db 50 IYIFIEFVIGMIANSVVVWNIQAKTTGYDTHCYILNLAIADLWVITIPVWVSVLQHNQ 109
QY 174 WVFGEWMCRIITVVFYGNMYCAILILTCMGINRYLATAHPFTYQKLPKRSFSLMCGIVM 233
Db 174 WVFGEWMCRIITVVFYGNMYCAILILTCMGINRYLATAHPFTYQKLPKRSFSLMCGIVM 233
QY 110 WPMGELTCKTHLIFSINLFGSIFFLACMSVDRLSITYFTSTSSYKKQVRRVVCVLVM 169
Db 110 WPMGELTCKTHLIFSINLFGSIFFLACMSVDRLSITYFTSTSSYKKQVRRVVCVLVM 169
QY 234 VMVFLYMLPFVILKQYHLVHSEITTCDDVDADACESPSFRFY-----YFVSL---A 282
Db 234 VMVFLYMLPFVILKQYHLVHSEITTCDDVDADACESPSFRFY-----YFVSL---A 282
QY 170 LLAF-----FVSLPDTYLL--KTVTSASNNETCYRS-----FYPEHSIKEWLIGMELVSV 217
Db 170 LLAF-----FVSLPDTYLL--KTVTSASNNETCYRS-----FYPEHSIKEWLIGMELVSV 217
QY 283 FFGFLIPFVIIIFCYTTLIHKLK-SKDRIWLGVIKAVLLILVIFTICFAPTNIIL----- 336
Db 283 FFGFLIPFVIIIFCYTTLIHKLK-SKDRIWLGVIKAVLLILVIFTICFAPTNIIL----- 336
QY 218 ILGFVAPFTIIAIFYELLARANSASGQEKHSRKIIFSVVVFLVCWLPYHFVLLDIF 277
Db 218 ILGFVAPFTIIAIFYELLARANSASGQEKHSRKIIFSVVVFLVCWLPYHFVLLDIF 277
QY 337 -VTHANYHYNTDSLYFMYLIALCLGSLNSCLDPPELYFVMSK 378
Db 337 -VTHANYHYNTDSLYFMYLIALCLGSLNSCLDPPELYFVMSK 378
QY 278 SILHYIPTCQENLVFLTAHLVTCQLSLVHCVCNVPVLYSFINR 320
Db 278 SILHYIPTCQENLVFLTAHLVTCQLSLVHCVCNVPVLYSFINR 320

RESULT 40
Q96HI6 PRELIMINARY; PRT; 331 AA.
AC Q96HI6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Unknown (protein for MGC:17339).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC008569; AA08569.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm1.1.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECP_F1_2; 1.
SQ SEQUENCE 331 AA; 38133 MW; 5B0B755A74FBB36 CRC64;

Query Match      14.2%; Score 303; DB 4; Length 331;
Best Local Similarity 30.2%; Pred. No. 6e-18;
Matches 83; Conservative 55; Mismatches 113; Indels 24; Gaps 11;

QY 116 YILLFVVGVPNIVTLWKLRLKTS-ISLVPHNTLAIDLFCVTLFPKIAVHLNGNN 174
Db 116 YILLFVVGVPNIVTLWKLRLKTS-ISLVPHNTLAIDLFCVTLFPKIAVHLNGNN 174
QY 29 YSCIFIIGLVNITALWVFSCTTKKTTVTIYMMNVALVDLFIIMLPFRMFYAK-DEW 87
Db 29 YSCIFIIGLVNITALWVFSCTTKKTTVTIYMMNVALVDLFIIMLPFRMFYAK-DEW 87
QY 175 VFGEVNCRI--TTVFYGNMYCAILILTCMGINRYLATAHPFTYQKLPKRSFSLMCGIV 232
Db 175 VFGEVNCRI--TTVFYGNMYCAILILTCMGINRYLATAHPFTYQKLPKRSFSLMCGIV 232

```

```

Db 88 PGEYFCQILGALTVPFYSI--ALWLLAFISADRYMAIVQPKYAKELKNTCKAVLACGV 145
QY 233 WMVFLYMLPFVILKQYHLVHSEITTCDDVDADACESPSFRFYFVSLAFTGFLIPFVI 292
Db 146 WIMTLTITTTPLLLLYKDDP-KDSTATCLKISDIYL-KAVNVLNLTTLTFF-FLIPLFI 202
QY 293 IIFCYTTLIHKL-----KSKDRIWLGVIKAVLLILVIFTICFAPTNI---ILVIHANY 343
Db 203 MIGCYLVIITHNLLHGRTSKLKPKEKSIRIIITLLVQLVLCFMPPHICFAFLMLGTGEN 262
QY 344 YHNTDSLYFMYLIALCLGSLNSCLDPPELYFVMSK 378
Db 263 SY-NPWGAPTTFM-----NLSTCLDVLIIYIVSK 291

Search completed: June 24, 2003, 12:05:35
Job time : 55.4368 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:59:21 ; Search time 25.9681 Seconds
(without alignments)
1506.722 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLTXXHPVAGSODIKKIL.....AMARPLXRRPRDIWDIHAH 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	582.5	27.3	420	2 I51667	thrombin receptor
2	581.5	27.2	399	2 I48705	proteinase activat
3	569	26.6	397	2 S66518	proteinase-activat
4	521	24.4	427	2 S17148	alpha-thrombin rec
5	512.5	24.0	432	2 A43448	thrombin receptor
6	509	23.8	425	2 A37912	thrombin receptor
7	431.5	20.2	361	2 B45680	G protein-coupled
8	399	18.7	362	2 S33733	G protein-coupled
9	389.5	18.2	370	2 JC5549	heptahelical P2Y5-
10	383	17.9	308	2 I50241	G protein-coupled
11	378.5	17.7	373	2 A47556	ATP receptor P2u -
12	376.5	17.6	373	2 JC4737	G protein-coupled
13	372	17.4	373	2 JC4162	P2Y receptor - bov
14	362	16.9	365	2 S68679	G protein-coupled
15	357	16.7	380	2 JC2434	kappa opioid recep
16	357	16.7	380	2 A48227	kappa opioid recep
17	356	16.7	344	2 T09508	intrin 17 purinerg
18	355.5	16.6	362	2 JN0694	angiotensin II rec
19	355	16.6	380	2 I38435	angiotensin recept
20	353.5	16.5	380	2 JC2338	kappa opioid recep
21	352	16.5	380	2 S36143	kappa opioid recep
22	350	16.4	380	2 A55259	kappa opioid recep
23	347.5	16.3	375	2 A54946	P-2U nucleotide re
24	333.5	15.6	362	2 B57641	G protein-coupled
25	333	15.6	328	2 I55450	G protein-coupled
26	332	15.5	359	2 I51372	angiotensin II rec
27	327	15.3	359	2 JC2134	angiotensin II rec
28	326	15.3	359	2 JC0621	angiotensin II rec
29	325	15.2	353	2 JC2492	G protein-coupled

30	324.5	15.2	372	2 S26667	G protein-coupled
31	324	15.2	359	2 JC1194	angiotensin II rec
32	321.5	15.1	354	2 I53033	G protein-coupled
33	321.5	15.1	362	2 A57641	G protein-coupled
34	320	15.0	359	2 S44425	angiotensin II rec
35	320	15.0	359	2 A48857	angiotensin II rec
36	320	15.0	359	2 S15403	angiotensin II rec
37	319.5	15.0	361	2 JC5653	G protein-coupled
38	318.5	14.9	365	2 S68208	G protein-coupled
39	318	14.9	358	2 A53752	interleukin-8 rece
40	317.5	14.9	328	2 JC4800	P2Y6 receptor - hu
41	317.5	14.9	341	2 S63666	platelet activatin
42	317.5	14.9	362	2 S68207	G protein-coupled
43	315	14.7	359	2 A42656	angiotensin II rec
44	314	14.7	355	2 A55733	G protein-coupled
45	314	14.7	355	2 A45177	chemokine (C-C) re

ALIGNMENTS

RESULT 1

I51667

Thrombin receptor - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C;Accession: I51667

R;Gerstzen, R.E.; Chen, J.; Ishii, M.;

Nature 368, 648-651, 1994

A;Title: Thrombin receptor's specificity for agonist peptide is determined by its extra

A;Reference number: I51667; MUID:94195429; PMID:8145852

A;Accession: I51667

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-420 <GER>

A;Cross-references: EMBL:U09632; NID:G495197; PIDN:AAA18498.1; PID:G495198

Query Match 27.3%; Score 582.5; DB 2; Length 420;

Best Local Similarity 32.7%; Pred. NO. 1.2e-38;

Matches 124; Conservative 82; Mismatches 138; Indels 35; Gaps 11;

QY	19	ILILVAAGLFLPVTVCQSGINVSNAK-----PTLTIKSFN--GGPQNTFEFFPLSDIE	72
DB	7	LLLLLLTLGLGMSLCLAN---SDTQAKGAHSNNMTIKTRIFDSESEFEIIPWDEL	63
QY	73	GMTGATTTIKAECPEDSISTL--HVNNTATIGVLRSSLSTQVIPAIYILLFVGVPSN---	127
DB	64	E-SGEGSQQAPVRSRARPDIRNITKEAEQVLSQWLTKFVPSLTVVFIUGLPLNLLA	122
QY	128	-IVTLWKLRLTKSISLVIFHTNLAIADLLFCVTLFPFKIAYHLNNGNWNVGEVYMCRTTV	186
DB	123	IIIFLPMKVRKPA---VVYMLNLAIADVFVSVLPFKIAYHLNGLNDLFGPGMCRIVTA	179
QY	187	VFGNNMYCAILLTCMGINRYLATAHPFYQKLPKRSPSLMCGIWMVWVLYMLPFVIL	246
DB	180	IFYCNNYCVLLIASISVDRLAVVYPMHSLSWRTMSRAYMACSFILWISIASTIPLLVT	239
QY	247	KQBYHLVHSEITTCDDVDVACSPSSFRFYVFLSAFFGLFPIFVLIIFCYTTLHLKLKS	306
DB	240	EQTKIPRIQITTCDDVLD-LKDLKDFIYIYFSSFCLLFFVFITTCIYIGIIRSLSS	298
QY	307	-----KDRILWGLYIKAVLLILVITFCFAPTNIILVIHANYYYHNTDSLYFYMLIA	358
DB	299	SSIENSCCKTRA---LFLAVVLCVFIICFGTNNVFLTH---YLQEANERFLYFAYILS	351
QY	359	LCGLNSCLDPFLYFVMS	377
DB	352	ACVGSVSCCLDPLIYYIAS	370

RESULT 2

I48705

proteinase activated receptor 2 - mouse

C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C;Accession: I48705
R;Nystedt, S.; Larsson, A.K.; Aberg, H.; Sundelin, J.
J. Biol. Chem. 270, 5950-5955, 1995
A;Title: The mouse proteinase-activated receptor-2 cDNA and gene. Molecular cloning and
A;Reference number: I48705; MUID:95197620; PMID:7890726
A;Accession: I48705
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-399 <RES>
A;Cross-references: EMBL:Z48043; NID:g663020; PIDN:CAA88097.1; PID:g663021
C;Superfamily: ATP receptor P2u

Query Match 27.2%; Score 581.5; DB 2; Length 399;
Best Local Similarity 37.6%; Pred. No. 1.3e-38;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY 95 VNNATIGYLRSSLSQVIPAIIYLLFVVGVPSSNIVTLWKLRLT-KSISLVIEHTNLATA 153
DB 63 IDEFSASILTGKLTVPVPIIVFVIGLPSNGMALWIFLTKKGPVAVIYMANLALA 122
QY 154 DLLFCVTLPPKIAHYHLGNWVGEVNCRTITVVYGNMYCAILILTCMGINRYLATAHP 213
DB 123 DLLSVIWFPLKISYHLHGNWVYGEALCKVLGFFYGNMYCSILFWTCLSVQRYWVIVNP 182
QY 214 FTQKLPKRSFSLMCGIVVMVFLYMLPFVILKQEVHLVHSEITTCDDVDVDAESPSSF 273
DB 183 MGHPK-KKANIAGVSAIWLIFLTPILYVMKQTIYIPALNITTCDDVLPK-EVLVGD 240
QY 274 RFYVFLAPFGFLIPPIVIFCVTTLI-----HKLKSKDRIMLGIYKAVILLIV 323
DB 241 MFNYFLSLATGVLPFALLTASAYVLMKILRSAMDEHSEKKRQA----IBLIITVLA 296
QY 324 IFTICFAPTNIIIVHANYHYNTDSLYFMYLIALCLGSLNSCLDPFLYFVMSK 378
DB 297 MYFICFAPSLLLVHVFYLIKTQKQSHVYALYVALCLSLNSCIDPPFVYFVSK 351

RESULT 3
S66518
proteinase-activated receptor 2 precursor - human
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Nov-2000
C;Accession: S66518; S64709; G02131
R;Nystedt, S.; Emission, K.; Larsson, A.K.; Stroembeck, B.; Sundelin, J.
Eur. J. Biochem. 232, 84-89, 1995
A;Title: Molecular cloning and functional expression of the gene encoding the human pro-
A;Reference number: S66518; MUID:96048032; PMID:7556175
A;Accession: S66518
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <NYS>
A;Cross-references: EMBL:Z49993; NID:g1008084; PIDN:CAA90290.1; PID:g1008085
R;Boehm, S.K.; Kong, W.; Broemme, D.; Smeekens, S.P.; Anderson, D.C.; Connolly, A.; Kahn
Biochem. J. 314, 1009-1016, 1996
A;Title: Molecular cloning, expression and potential functions of the human proteinase-a
A;Reference number: S64709; MUID:96177879; PMID:8615752
A;Accession: S64709
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-137, 'A', 139-397 <BOE>
A;Cross-references: EMBL:U34038; NID:g1041728; PIDN:ABA47871.1; PID:g1041729
A;Note: the authors translated the codon GTC for residue 68 as Ile and AAC for residue 2
R;Kahn, M.L.; Coughlin, S.R.
submitted to the EMBL Data Library, September 1995
A;Reference number: H00822
A;Accession: G02131
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 29-397 <KAH>
A;Cross-references: EMBL:U36753; NID:g1208539; PIDN:AAA90957.1; PID:g1208540
C;Genetics:

A;Map position: 5q13
A;Introns: 28/1
C;Superfamily: ATP receptor P2u
F;1-36/Domain: activation peptide #status predicted <APT>
F;1-25/Domain: signal sequence #status predicted <SIG>
F;37-397/Product: proteinase-activated receptor 2 #status predicted <MAT>

Query Match 26.6%; Score 569; DB 2; Length 397;
Best Local Similarity 32.7%; Pred. No. 1.3e-37;
Matches 120; Conservative 73; Mismatches 136; Indels 38; Gaps 8;

QY 22 LVAAGLLFLPVVTCQSGINVSNSAKPTLTIKSFNGGPONTPEEPPLSDIEGWTGATTTI 81
DB 9 LIGAALLAASLSCSGTIOGTNRSSKGRSLICKVDG-----TSHV---TCKGVTV 55
QY 82 KAECPEDSISTLHVNNATIGYLRSSLSQVIPAIIYLLFVVGVPSSNIVTLWKLRLT-KS 140
DB 56 E-----TFVSDFEFSASVLTGKLTVPVPIIVFVIGLPSNGMALWIFLTKK 107
QY 141 ISLVIFHTNLATADLLFCVTLPPKIAHYHLGNWVGEVNCRTITVVYGNMYCAILILT 200
DB 108 HPVAVIMANLADLLSVIWFPLKIAHYHLHGNWYGEALCNVLIGFFYGNMYCSILFWT 167
QY 201 CMGINRYLATAHPFTYQKLPKRSFSLMCGIVVMVFLYMLPFVILKQEVHLVHSEITTC 260
DB 168 CLSVQRYWIVNPMGHSR-KKANIAIGISLAIWLLILVITPLVYVMKQTIYIPALNITTC 226
QY 261 HDVDACESPSSF-----RFYVFLAPFGFLIPFVILKQEVHLVHSEITTC 310
DB 227 HDVL-----PEQLLYGDMFNFLSLAIGVFLPFAFLTASAYVLMIRMLRSSAMDEHSEKK 281
QY 311 WLGYIKAVILLIVIFTCFAPTNIIIVHANYHYNTDSLYFMYLIALCLGSLNSCLDP 370
DB 282 RRKRAKLIVTLVAMLICTPSNLLLVHVFYLIKTQKQSHVYALYVALCLSLNSCIDP 341
QY 371 FLYFVMS 377
DB 342 FVYFVS 348

RESULT 4
S17148
alpha-thrombin receptor - Chinese hamster
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C;Accession: S17148
R;Rasmussen, U.B.; Vouret-Craviari, V.; Jallat, S.; Schlesinger, Y.; Pages, G.; Paviran
FEBS Lett. 288, 123-128, 1991
A;Title: cDNA cloning and expression of a hamster alpha-thrombin receptor coupled to Ca
A;Reference number: S17148; MUID:91348247; PMID:1652467
A;Accession: S17148
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-427 <RAS>
A;Cross-references: EMBL:X61958; NID:g940495; PIDN:CAA43957.1; PID:g49538
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 24.4%; Score 521; DB 2; Length 427;
Best Local Similarity 32.4%; Pred. No. 8.7e-34;
Matches 129; Conservative 78; Mismatches 125; Indels 66; Gaps 16;

QY 20 LILVAAG-----LLFLPVVTCQSGINVSNSAKPTLTIKSF---NGGPONTPEEPPLSD 70
DB 6 LLLVAAAGSLCGLPLSSRVVVPVQPESEMDATVNP-----RSFRLNPG-ENTFELIPLGD 60
QY 71 IEGWTGATTTIKAECPEDSISTLH-----VNNATIGYLRSSLSQVIPAIIYLL 119
DB 61 EEKNEST-----LPEGRAIYLNKSHSPALPFIFSEDASGYLTSPLWRLRFPSPVTVFV 114
QY 120 FVVGVPSSNI--VTLWKLRLT-KSISLVIEHTNLATADLLFCVTLPPKIAHYHLGNWVFG 177
DB 115 FVVSPLNLTAIAVFLVKMKVKK-PAVVYMLHLAMADVLSVLPLKISYFSGSDMQFG 173

A:Gene: p2Y5
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; transmembrane protein
F:15-40/Domain: transmembrane #status predicted <TM1>
F:51-74/Domain: transmembrane #status predicted <TM2>
F:89-109/Domain: transmembrane #status predicted <TM3>
F:133-153/Domain: transmembrane #status predicted <TM4>
F:177-201/Domain: transmembrane #status predicted <TM5>
F:227-248/Domain: transmembrane #status predicted <TM6>
F:269-292/Domain: transmembrane #status predicted <TM7>

Query Match 17.9%; Score 383; DB 2; Length 308;
Best Local Similarity 30.3%; Pred. No. 5.2e-23;
Matches 90; Conservative 59; Mismatches 104; Indels 44; Gaps 9;

QY 115 IYILFVVGVPNSIVTLW---KLSLRKTSISLVIFHTNLAIADLLFCVTLFPFKIAYHLN 170
DB 20 VFSMVFLGLIANCAVIAIYFTTLKVRNETTYML---NLASDLLFVTLFRI-YYPV 75

QY 171 GNNVFEVGMCRITTVVYGNMYCAIILTCMGIRNYLATAHPFTYQKLPKRSFSLMCG 230
DB 76 VRNWPFGDVLCKISVTLFTYNNYGSILFTLCISVDRLFAIVHFRSKTLRTKRNARIVCV 135

QY 231 IYVWVFLYMLFPVILKQYHLVHSEITTCHEVDVADACSPSPFRFYFVSLAFF----CF 286
DB 136 AVMITVLAGSTPASFFQSTNRQNNTEQRTCFE-----NFPSTWTKYLSRIVIFTEIVGF 190

QY 287 LIPFIIIFCYTTLIHKLK-----SKRIWLGVIKAVILLIVIFTICFAPTNIL 336
DB 191 FPIILNVTCSTWLTNLKPLTLNRKLSKKV-----LKMFLVHLVIFCFVFPNITL 246

QY 337 VIHANYYYHNTD-----SLYPMYLIALCLGSLNSCLDPFLYFVMSKVVDQLN 384
DB 247 IL----YSLMRIQTWNCVWTAVRTWYPTVLCIAVSNCCFDPVIVYFTSDTNSELD 299

RESULT 11
ATP receptor P2u - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C:Accession: A47556
R:Jusztig, K.D.; Shiau, A.K.; Brake, A.J.; Julius, D.
Proc. Natl. Acad. Sci. U.S.A. 90, 5113-5117, 1993
A:Title: Expression cloning of an ATP receptor from mouse neuroblastoma cells.
A:Reference number: A47556; MUID:93281707; PMID:7685114
A:Accession: A47556
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-373 <LUS>
A:Cross-references: GB:L14751; NID:G309457; PIDN:AAA39871.1; PID:G309458
C:Superfamily: ATP receptor P2u
C:Keywords: transmembrane protein

Query Match 17.7%; Score 378.5; DB 2; Length 373;
Best Local Similarity 26.9%; Pred. No. 1.4e-22;
Matches 100; Conservative 67; Mismatches 144; Indels 61; Gaps 13;

QY 69 SDIEGTGATTIKACRPEDSISTLHVNNATIGYLRSSLSQTQVIPAIIYLLFVVGVPNI 128
DB 3 ADLEPW---NSTINGTWEGDELGYKCRFNEDFKYV-----LLPVSYGWCVLGLCLNV 52

QY 129 VTIWKLRLTK-----SISLVIFHTNLAIADLLFCVTLFPFKIAYHLNNGVGMCRIT 185
DB 53 VALYIFLCRLKTNASTTYWFH---LAVSDSLYASLPLIVYYARGDHPFFSTVLCKLYR 110

QY 186 VFYGNMYCAIILTCMGIRNYLATAHPFTYQKLPKRSFSLMCGIYVWVFLYMLP--- 242
DB 111 FLPTNLYCSILFTLCISVHRCGLVRLPLSLRWGRARYARRVAVWVVLVACQAPVLY 170

QY 243 FVILKQYHLVHSEITTCHEVDVADACSPSPFRFYFVSLAFFGLIFPVIIIFCYTTLIH 302
DB 171 FVTTS-----VRGTRITCHD-TSARELFHFVAYSSVMLGLL-FAVPFPSVILVYCYVLMAR 223

QY 303 KL-----KSKRIWLGVIKAVILLIVIFTICFAPTNILVIHHA-----NYY 344
DB 224 RLLKPAYGTGGLPRAKRS-----VRTIALVAVFALCFLPFHVRTLYYSFRSLDLS 277

QY 345 YHNTDSLFPMYLIALCLGSLNSCLDPFLYFVMSKV-----DQINPXSNMARP----- 392
DB 278 CHTLNAINMAYKITRPLASANSCLDPVLYFLAGQLRVFARDAKPPTEPTSPQARRKLG 337
QY 393 LXRPDRDIWEDI 404
DB 338 LHRPNRTVRKDL 349

RESULT 12
JC4737
G protein-coupled receptor P2Y1 - human
N:Alternate names: P2Y1 purinergic receptor; P2Y1 purinoceptor
C:Species: Homo sapiens (man)
C:Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 17-Nov-2000
C:Accession: JC4737; JC4615; S54253
R:Janssens, R.; Communi, D.; Pirotton, S.; Samson, M.; Parmentier, M.; Boeynaems, J.M.
Biochem. Biophys. Res. Commun. 221, 588-593, 1996
A:Title: Cloning and tissue distribution of the human P2Y1 receptor.
A:Reference number: JC4737; MUID:96205320; PMID:8630005
A:Accession: JC4737
A:Molecule type: DNA
A:Residues: 1-373 <GAN>
A:Cross-references: GB:S81950; NID:G1839438; PIDN:AA847091.1; PID:G1839439
R:Ayyanathan, K.; Webb, T.E.; Sandhu, A.K.; Athwal, R.S.; Barnard, E.A.; Kunapuli, S.P.
Biochem. Biophys. Res. Commun. 218, 783-788, 1996
A:Title: Cloning and chromosomal localization of the human P2Y1 purinoceptor.
A:Reference number: JC4615; MUID:96158962; PMID:8579591
A:Accession: JC4615
A:Molecule type: mRNA
A:Residues: 1-373 <AYY>
A:Cross-references: GB:U42029; NID:G1147730; PIDN:AAA97872.1; PID:G1147731
A:Experimental source: erythro leukemia cells
R:Leon, C.; Vial, C.; Cazenave, J.; Gachet, C.
submitted to the EMBL Data Library, May 1995
A:Description: Cloning of a human putative P2Y receptor.
A:Reference number: S54253
A:Accession: S54253
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-137,139-373 <LEO>
A:Cross-references: EMBL:Z49205; NID:G798835; PIDN:CAA83066.1; PID:G798836
C:Comment: This receptor belongs to a family of G protein-coupled receptors. It responds:
C:Genetics:
A:Gene: p2Y1; GDB:P2RY1
A:Cross-references: GDB:677125; OMIM:601167
A:Map position: 3pter-3qter
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane pro.
F:52-77/Domain: transmembrane #status predicted <TM1>
F:88-111/Domain: transmembrane #status predicted <TM2>
F:124-152/Domain: transmembrane #status predicted <TM3>
F:171-191/Domain: transmembrane #status predicted <TM4>
F:214-237/Domain: transmembrane #status predicted <TM5>
F:261-282/Domain: transmembrane #status predicted <TM6>
F:305-328/Domain: transmembrane #status predicted <TM7>
F:11,27,113,197/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:258,336/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predic
F:330,339/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predic
F:343/Binding site: phosphate (Ser) (covalent) (by protein kinase C and calmodulin-depe

Query Match 17.6%; Score 376.5; DB 2; Length 373;
Best Local Similarity 27.0%; Pred. No. 2.1e-22;
Matches 95; Conservative 72; Mismatches 130; Indels 55; Gaps 12;

QY 46 AKPTLTIKSFGNGPONTPEEPFLSDIEGTGATTTIKACRPEDSISTLHVNNNA-TIGYLR 104
DB 8 AVPNGTDAFLAGPGSS-----WGNSTVA-----STAASVSSFKCALTK 46

Query Match 16.7%; Score 357; DB 2; Length 380;
Best Local Similarity 27.9%; Pred. No. 7.4e-21;
Matches 100; Conservative 66; Mismatches 124; Indels 68; Gaps 12;

QY 52 IKSFNGGPQNT-----FEERPLSDIEGTGATTITIKAECPEDSISTLHVNN 97
DB 5 IQIFRGDPGPTCPSACLLPNSSWFPNWAESDNGSVGS-----EDQLESASHSP 56
QY 98 ATIGYLRSSLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLTK---SISLVIFHTNLAIAD 154
DB 57 AI-----PVIITAVYSVWFVGLVGNLSLWVFVIIRYTKMKTATNIYIF--NLALAD 105
QY 155 LLFCVTLPPKIAVHLNGNNWVGEVMCRITTVVFGNYCAILITCMGINRYLATAHPF 214
DB 106 ALVTTMTFQSAVYLM--NSWPFGLVKLVISIDYNNMFTSIFTLTMMSDRYIAVCHPV 164
QY 215 TYQ--KLKRSFSLMCGVWVMFLYMLPFVILKQEHVHSEITTHCHDVVDACESPSS 272
DB 165 KALDFRTPLKAKIINIC--IWLASSVGISAIVLGG-----TKVREDVDVIECLQ 213
QY 273 F-----RFYFVSLAFFGLIPFVIIIFCYTTLIHKLKS-----KDRIMLGYI 315
DB 214 FPDDEYSWMDLPMKICVVFAPVILVIVCYTTLMLRLKSVRLLSGSGREKDRNLRRIT 273
QY 316 KAVLLILVITFCFAPTNIILVIHANYVYHNTDSLYFMYLIALCLGSLNSCLDPLFY 373
DB 274 KLVVVAVFIICTWPIHIFILVEALGSTSHSTAAL--SSYFICIALGYTNSSINPVLV 330
RESULT 16
A48227
kappa opioid receptor 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jan-2000
C:Accession: A48227; J04138
R:Yasuda, K.; Raynor, K.; Kong, H.; Breder, C.D.; Takeda, J.; Reisine, T.; Bell, G.I.
Proc. Natl. Acad. Sci. U.S.A. 90, 6736-6740, 1993
A:Title: Cloning and functional comparison of kappa and delta opioid receptors from mouse
A:Reference number: A48227; MUID:93342064; PMID:8393575
A:Accession: A48227
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-380 <YAS>
A:Cross-references: GB:I11065; NID:G348248; PIDN:AAA39363.1; PID:G348249
R:Liu, H.C.; Lu, S.; Augustin, L.B.; Felsheim, R.F.; Chen, H.C.; Loh, H.H.; Wei, L.N.
Biochem. Biophys. Res. Commun. 209, 639-647, 1995
A:Title: Cloning and promoter mapping of mouse kappa opioid receptor gene.
A:Reference number: J04138; MUID:95251663; PMID:7733933
A:Accession: J04138
A:Molecule type: mRNA
A:Residues: 1-380 <LIU>
A:Note: The authors translated the codon CAG for residue 365 as Glu
C:Comment: This receptor exists in different areas of the central and peripheral nervous
C:Genetics:
A:Gene: kor
C:Superfamily: vertebrate rhodopsin
C:Keywords: brain; G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane

Query Match 16.7%; Score 357; DB 2; Length 380;
Best Local Similarity 27.9%; Pred. No. 7.4e-21;
Matches 100; Conservative 66; Mismatches 124; Indels 68; Gaps 12;

QY 52 IKSFNGGPQNT-----FEERPLSDIEGTGATTITIKAECPEDSISTLHVNN 97
DB 5 IQIFRGDPGPTCPSACLLPNSSWFPNWAESDNGSVGS-----EDQLESASHSP 56
QY 98 ATIGYLRSSLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLTK---SISLVIFHTNLAIAD 154
DB 57 AI-----PVIITAVYSVWFVGLVGNLSLWVFVIIRYTKMKTATNIYIF--NLALAD 105
QY 155 LLFCVTLPPKIAVHLNGNNWVGEVMCRITTVVFGNYCAILITCMGINRYLATAHPF 214
DB 106 ALVTTMTFQSAVYLM--NSWPFGLVKLVISIDYNNMFTSIFTLTMMSDRYIAVCHPV 164
QY 215 TYQ--KLKRSFSLMCGVWVMFLYMLPFVILKQEHVHSEITTHCHDVVDACESPSS 272
DB 165 KALDFRTPLKAKIINIC--IWLASSVGISAIVLGG-----TKVREDVDVIECLQ 213
QY 273 F-----RFYFVSLAFFGLIPFVIIIFCYTTLIHKLKS-----KDRIMLGYI 315
DB 214 FPDDEYSWMDLPMKICVVFAPVILVIVCYTTLMLRLKSVRLLSGSGREKDRNLRRIT 273
QY 316 KAVLLILVITFCFAPTNIILVIHANYVYHNTDSLYFMYLIALCLGSLNSCLDPLFY 373
DB 274 KLVVVAVFIICTWPIHIFILVEALGSTSHSTAAL--SSYFICIALGYTNSSINPVLV 330
RESULT 16
A48227
kappa opioid receptor 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jan-2000
C:Accession: A48227; J04138
R:Yasuda, K.; Raynor, K.; Kong, H.; Breder, C.D.; Takeda, J.; Reisine, T.; Bell, G.I.
Proc. Natl. Acad. Sci. U.S.A. 90, 6736-6740, 1993
A:Title: Cloning and functional comparison of kappa and delta opioid receptors from mouse
A:Reference number: A48227; MUID:93342064; PMID:8393575
A:Accession: A48227
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-380 <YAS>
A:Cross-references: GB:I11065; NID:G348248; PIDN:AAA39363.1; PID:G348249
R:Liu, H.C.; Lu, S.; Augustin, L.B.; Felsheim, R.F.; Chen, H.C.; Loh, H.H.; Wei, L.N.
Biochem. Biophys. Res. Commun. 209, 639-647, 1995
A:Title: Cloning and promoter mapping of mouse kappa opioid receptor gene.
A:Reference number: J04138; MUID:95251663; PMID:7733933
A:Accession: J04138
A:Molecule type: mRNA
A:Residues: 1-380 <LIU>
A:Note: The authors translated the codon CAG for residue 365 as Glu
C:Comment: This receptor exists in different areas of the central and peripheral nervous
C:Genetics:
A:Gene: kor
C:Superfamily: vertebrate rhodopsin
C:Keywords: brain; G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane

DB 106 ALVTTMTFQSAVYLM--NSWPFGLVKLVISIDYNNMFTSIFTLTMMSDRYIAVCHPV 164
QY 215 TYQ--KLKRSFSLMCGVWVMFLYMLPFVILKQEHVHSEITTHCHDVVDACESPSS 272
DB 165 KALDFRTPLKAKIINIC--IWLASSVGISAIVLGG-----TKVREDVDVIECLQ 213
QY 273 F-----RFYFVSLAFFGLIPFVIIIFCYTTLIHKLKS-----KDRIMLGYI 315
DB 214 FPDDEYSWMDLPMKICVVFAPVILVIVCYTTLMLRLKSVRLLSGSGREKDRNLRRIT 273
QY 316 KAVLLILVITFCFAPTNIILVIHANYVYHNTDSLYFMYLIALCLGSLNSCLDPLFY 373
DB 274 KLVVVAVFIICTWPIHIFILVEALGSTSHSTAAL--SSYFICIALGYTNSSINPVLV 330
RESULT 17
T09508
intron 17 purinergic receptor P2Y5 - human
N:Alternate names: G-protein coupled receptor
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C:Accession: T09508
R:Bohm, S.K.; Trump, A.; Khitin, L.M.; Kong, W.; Pavan, D.G.; Bunnett, N.W.
submitted to the EMBL Data Library, April 1997
A:Description: The human purinergic receptor P2Y5 is encoded in intron 17 of the retinol
A:Reference number: Z16705
A:Accession: T09508
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-344 <BOH>
A:Cross-references: EMBL:AF000546; NID:G22322068; PID:G22322069
C:Genetics:
A:Map position: 13
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.7%; Score 356; DB 2; Length 344;
Best Local Similarity 28.2%; Pred. No. 8e-21;
Matches 87; Conservative 72; Mismatches 110; Indels 40; Gaps 11;

QY 97 NATIGYLRSSLSTQVIPAIIYLLFVVGVPNSIVTLW----KLSLRKTSISLVIFHTNLAI 152
DB 5 NSSHCIFYNDSFKYTYGCMFSWVFLGLVSNCAIYIFICVLKVRNETTYMI---NLAM 61
QY 153 ADLLFCVTLPPKIAVHLNGNNWVGEVMCRITTVVFGNYCAILITCMGINRYLATAH 212
DB 62 SDLLFVTLPPRI--FYFTRNWPFGDLLCKISVLMFYTNMYSILFLTCISVDRFLAIVY 120
QY 213 PTYQKLPKRSFSLMCGVWVMFLYMLPFVILKQEHVHSEITTHCHDVVDAC--ESP 270
DB 121 PFKSKTLRTKNAKIVCTGVMLTVIGGSAPAFVQS---THSQ---GNNASEACFENFP 173
QY 271 SFPRFYVPSLAFF----GFLIPFVIIIFCYTTLIHKLK-----SKDRIMLGYIK 316
DB 174 EATWKYLSRIVIFIEIVGFFIPLINLVTCSSMVLKTLTKPVTLSRSKINKTKV----LK 229
QY 317 AVLLILVITFCFAPTNIILVHH--ANYVYHNTD---SLYPMYLIALCLGSLNSCLDPPF 371
DB 230 MIFVHLIIFCFVCPVYNINILYSLVTRQTFVNCVAAVRYMYPITLCIAVSNCCFDPI 289
QY 372 LYFVMSKVY 380
DB 290 VYVFTSDTI 298
RESULT 18
JN0694
angiotensin II receptor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: JN0694; S47627
R:Ji, H.; Sandberg, K.; Zhang, Y.; Catt, K.J.
Biochem. Biophys. Res. Commun. 194, 756-762, 1993

A:Title: Molecular cloning, sequencing and functional expression of an amphibian angiotensin receptor
A:Reference number: JN0694; MUID:93343933; PMID:7688227
A:Accession: JN0694
A:Molecule type: mRNA
A:Residues: 1-362 <JH>
A:Cross-references: GB:116463; NID:9387890; PIDN:AAA49647.1; PID:9387891
R:Nishimatsu, S.; Koyasu, N.; Sugaya, T.; Ohnishi, J.; Yamagishi, T.; Murakami, K.; Miya Biochim. Biophys. Acta 1218, 401-407, 1994
A:Title: Isolation and characterization of two alternatively spliced complementary DNAs
A:Reference number: S47627; MUID:94325348; PMID:7519446
A:Accession: S47627
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-362 <NIS>
A:Cross-references: GB:S73734; NID:G625009; PIDN:AAC60749.1; PID:G625010
C:Superfamily: vertebrate rhodopsin
C:Keywords: glycoprotein; phosphoprotein; transmembrane protein
F:34-54/Domain: transmembrane #status predicted <TM1>
F:66-86/Domain: transmembrane #status predicted <TM2>
F:105-125/Domain: transmembrane #status predicted <TM3>
F:146-166/Domain: transmembrane #status predicted <TM4>
F:192-212/Domain: transmembrane #status predicted <TM5>
F:241-260/Domain: transmembrane #status predicted <TM6>
F:285-305/Domain: transmembrane #status predicted <TM7>
F:318-177/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:322,343/Binding site: phosphate (Thr) (covalent) #status predicted
F:347/Binding site: phosphate (Ser) (covalent) #status predicted
Query Match 16.6%; Score 355.5; DB 2; Length 362;
Best Local Similarity 29.1%; Pred. No. 9.2e-21;
Matches 92; Conservative 56; Mismatches 127; Indels 41; Gaps 9;
QY 89 SISTLVHNNATIGYLRSSLSSTQV---IPATYILLFVVGVPNS--IVTLWKLISRTKSL 143
DB 6 TVETSDVERIANVCSGSHNYIFIAPIIYSTIFVVGFGNSMVVIYISYNNKMTVA- 64
QY 144 VIFHTNLAIADLLFCVLPFKIAYHLNGNNWVGEVNCRIITVVFYGNMYCAILLITCMG 203
DB 65 SIFLMNLASDLFCVITLPLWAAYTAMHYHWPFGNFCVKVASTAITNLNLTIVTFLITCL 124
QY 204 INRYLATAPFTYQKLPKRSFSLMCGIYVWVFLYMLPFVILKQYHLVHSEITTCYD 263
DB 125 IDRYSAIVHPMKSRIRWTAMVARLTCVGIWLAFLASPSIIRQIYLEFHDNTQVCAIV 184
QY 264 VDACESPSSRFYFVSLA----FFGLIPFVILIFCYTTLIHLKAS-----KDRIW 311
DB 185 YD-----SGHIYFMVGMGLAKNIVGFLIPFLIILTSYTLGKTLKEVYRAQRNDDIF 238
QY 312 LGYIKAVLLILVITTCFAPTNIIL-----VIHANYYYHNTDSLYFMYLIACLG 362
DB 239 ----KMIVAVLLFFFCWIPYQVFTFLDVLIQMDVIONCKWY----DIVDTGMPITICIA 290
QY 363 SLNSCLDPLFYFVNSK 378
DB 291 YFNSCLNPLFYGFPGK 306
RESULT 19
I38435
angiotensin receptor homolog APJ - human
C:Species: Homo sapiens (man)
C:Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 21-Jul-2000
C:Accession: I38435
R:O'Dowd, B.F.; Heiber, M.; Chan, A.; Heng, H.H.; Tsui, L.; Kennedy, J.L.; Shi, X.; Petr Gene 136, 355-360, 1993
A:Title: A human gene that shows identity with the gene encoding the angiotensin receptor
A:Reference number: I38435; MUID:94124031; PMID:8294032
A:Accession: I38435
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-380 <RES>
A:Cross-references: EMBL:U03642; NID:9425351; PIDN:AAA18954.1; PID:9425352
C:Genetics:

A:Gene: APJ
A:Map position: 11q12
A:Introns: #status absent
C:Superfamily: vertebrate rhodopsin
Query Match 16.6%; Score 355; DB 2; Length 380;
Best Local Similarity 27.3%; Pred. No. 1.1e-20;
Matches 89; Conservative 66; Mismatches 127; Indels 44; Gaps 9;
QY 70 DIEGTGATTIKACPCPDSISTLVHNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIV 129
DB 6 DFDNYTGADN--QSECE-----YTDKSSGALIPAIYMLVFLGTGTGNGL 48
QY 130 TLWKL--SLRTKISLVIFHTNLAIADLLFCVLPFKIAYHLNGNNWVGEVNCRIITTV 187
DB 49 VLVTVFSSREKRSRADIPIASLAVADLTFTVTLPLWATYTYRDYDNFEGTFCKLSVL 108
QY 188 FYGNMYCAILLITCMGINRYLATAHPFTYQKLPKRSFSLMCGIYVWVFLYMLPFVILK 247
DB 109 IFVNNYASVFCITGLSDFRYLAIVRPVANARLRLRVSGAVATAVLWLAALLAMPVWL 168
QY 248 QEYHLVHSEITTC---HDVDACESPSSRFYFVSLAFFGLIPFVILIFCYTTL---- 300
DB 169 TTGDLNTTKVQCYNDYSWATVSEWAVEGLVSSITVGVFVPTIMLTCTCYFFIAQTI 228
QY 301 -----IHLKSKDRIWLGVIKAVLLILVITTCFAPTNIILVIHANYYYH---NTD 349
DB 229 AGHPRKERIEGLRKRRRL-----LSIIVLVITFALCWMPYHLVLYMLGSLHWP 284
QY 350 SLYFMYLI--ALCLGSLNSCLDPLFY 373
DB 285 -LFLMNIFFPYCTCISYVNSCLNPLFY 309
RESULT 20
JC2338
kappa opioid receptor 1 - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 19-May-2000
C:Accession: JC2338; A55354; I57005; G01546
R:Mansson, E.; Bare, L.; Yang, D. Biochem. Biophys. Res. Commun. 202, 1431-1437, 1994
A:Title: Isolation of a human kappa opioid receptor cDNA from placenta.
A:Reference number: JC2338; MUID:94338360; PMID:8060324
A:Accession: JC2338
A:Molecule type: mRNA
A:Residues: 1-380 <MAN>
A:Cross-references: GB:U11053; NID:9532059; PIDN:AAA20985.1; PID:9532060
A:Experimental source: placenta
R:Wang, J.B.; Johnson, P.S.; Wu, J.M.; Wang, W.F.; Uhl, G.R. J. Biol. Chem. 269, 25966-25969, 1994
A:Title: Human kappa opiate receptor second extracellular loop elevates dynorphin's affinity
A:Reference number: A55354; MUID:95014415; PMID:7529306
A:Accession: A55354
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 136-279 <WAN>
A:Cross-references: GB:L36130; NID:9598184; PIDN:AAA63646.1; PID:9598185
R:Zhu, J.; Chen, C.; Xue, J. Life Sci. 56, 201-207, 1995
A:Title: Cloning of a human kappa opioid receptor from the brain.
A:Reference number: I57005
A:Accession: I57005
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1,'D',3-380 <ZHU>
R:Grandy, D.K. submitted to the EMBL Data Library, November 1994
A:Reference number: G07718
A:Accession: G01546
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA

Query Match 16.4%; Score 350; DB 2; Length 380;
Best Local Similarity 27.6%; Pred. No. 2.6e-20;
Matches 96; Conservative 65; Mismatches 131; Indels 56; Gaps 11;

QY 58 GQNTFEERPLSD-----IEGWTGATTIKACPEDSISILHVNNAIGYLRSL 107
DB 7 GPAQASELPARNACLLPNAGSAPDNGSAGPODELEPAHISPAI-----58

QY 108 STQVPAIYLLFVVGVPNSIVTLWKLSTK--SISLVIFHTNLAIAADLLFCVTLPPK 164
DB 59 -PVIITAVSVFVVGVLGNSLVMFVIIRYTKMKATNIYIF--NLAJADALVTTMPQ 115

QY 165 IAYHLNGNNWFGVCMCRITTVVYGNMYCAIILITCMGINRYLATAHPTTYQ--KLPRK 222
DB 116 STVYLM-NSWPGDVLCKIVISIDYNNMFTSIFLTMSVDVRYIAVCHPVKALDFRTPLK 174

QY 223 SFSLLMCGIWMVFLYMLPFVILKQEHVHSEITTCDDVDVACESPSSF-----R 274
DB 175 AKIINIC--IWLSSSVGISAILGG-----TKREDVDIIECSLQFPDDDDYSWMD 223

QY 275 FYFVSLAFEGFLIPVILIFCVTLIHLKLS-----KDRWLGYIKAVLLILVIF 325
DB 224 LFMKICVFVAFVPIVILIIIVCYTLMLKLSVRLSSGREGKDRNLRRITRLVLVVAVF 283

QY 326 TICFAPTNILVIHANYYYHNTDSLYFMYLIACLSLNSCLDPPLY 373
DB 284 IICWTPIHIFILVEALGSTSHSTAAL--SSYFICIALGYTNSSINPLY 330

RESULT 23
A54946
P-2U nucleotide receptor - human
C:Species: Homo sapiens (man)
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 17-Mar-1999
C:Accession: A54946
R:Parf, C.E.; Sullivan, D.M.; Paradiso, A.M.; Lazarowski, E.R.; Burch, L.H.; Olsen, J.C.
Proc. Natl. Acad. Sci. U.S.A. 91, 3275-3279, 1994
A:Title: Cloning and expression of a human P-2U nucleotide receptor, a target for cystic
A:Reference number: A54946; MUID:94211846; PMID:8159738
A:Accession: A54946
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-375 <PAR>
A:Cross-references: GB:U07225
A>Note: parts of this sequence were confirmed by protein sequencing
C:Genetics:
A:Gene: GDB:P2R2; HP2U; P2U
A:Cross-references: GDB:362713; OMIM:600041
A:Map position: 11q13.5-11q14.1
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.3%; Score 347.5; DB 2; Length 375;
Best Local Similarity 28.7%; Pred. No. 4.1e-20;
Matches 83; Conservative 57; Mismatches 108; Indels 41; Gaps 10;

QY 111 VIPAIYLLFVVGVPNSIVTLWKLSTK--SISLVIFHTNLAIAADLLFCVTLPEKIAY 167
DB 35 LLPVSVGVVGVGLCLNLAVGLYIFLCRLKTNASTYMFH--LAVSDALYAAASLLPLVY 92

QY 168 HUNGNNWFGVCMCRITTVVYGNMYCAIILITCMGINRYLATAHPTTYQKLPKRSFSL 227
DB 93 YARGDHWPFSTVLCKLVRELFTYNLCSILFTCSVHRCGLVRLSLRWGRARYAR 152

QY 228 MCGIWMVFLYMLP---FVILKQEHVHSEITTCDDVDVACESPSSFRFYFVSLAPF 284
DB 153 VAGAVVWLVLACQAPVLYFVTTISARGPL-----TCHD-TSAPELFSRFVAYSSVMLGL 205

QY 285 GELIPVILIFCVTLIHLK-----KSKDRWLGYIKAVLLILVIFITCFA 330
DB 206 -FAVPFVILVUCVLMARLLKPAYTSGGLPRAKRS-----VRTIAVLAVFALCF 258

QY 331 PTNIIILVIHHA-----NYYHNTDSLYFMYLIACLSLNSCLDPPLYFV 375

Db 259 PFHVTRTLYYSFRSLDLSCHTLNAINMAYKVR-LASANSCLDPVLYEL 306
RESULT 24
B57641
G protein-coupled receptor 4 - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 29-Sep-1999
C:Accession: B57641
R:Mahadevan, M.S.; Baird, S.; Bailly, J.E.; Shutler, G.G.; Sabourin, L.A.; Tsilfidis, C.
Genomics 30, 84-88, 1995
A:Title: Isolation of a novel G protein-coupled receptor (GPR4) localized to chromosome
A:Reference number: A57641; MUID:96129306; PMID:8595909
A:Accession: B57641
A>Status: preliminary; nucleic acid sequence not shown; translation not shown; significance
A:Molecule type: DNA
A:Residues: 1-362 <MAH>
A:Cross-references: GB:U22108; NID:g722282; PIDN:AAA98458.1; PID:g722283
C:Genetics:
A:Introns: #status absent
C:Superfamily: G protein-coupled receptor 4
C:Keywords: G protein-coupled receptor

Query Match 15.6%; Score 333.5; DB 2; Length 362;
Best Local Similarity 28.2%; Pred. No. 5.1e-19;
Matches 81; Conservative 60; Mismatches 109; Indels 37; Gaps 7;

QY 113 PAIYILLFVVGVPNSIVTLWKL--SLRTKSTSLVIFHTNLAIAADLLFCVTLPEKIAYHLN 170
DB 21 PSLYIFVIGVGLPTNCLRLMAAYRQVRN--ELGVYLMNLSTADLLYICTLPLWVDYFLH 79

QY 171 GNNVVFGEVCMCRITTVVYGNMYCAIILITCMGINRYLATAHPTTYQKLPKRSFSLMCG 230
DB 80 HDNHTHGPGCKLFGFIFTYIYISIAFLCCISVDVRYLAHAHLRFAHLRVKTAVAVSS 139

QY 231 IWMVFLYMLPFLVILKQEHVHSEITTCDDVDVACESPSSFRFYF-----VS 280
DB 140 VVMAT-----ELGANSVPLFDELFRDYNHTFCFEKFPMEGVAMWNL 184

QY 281 LAFFGFLIPVILIFCVTLIHLK---SKDRWLGYIKAVLLILV-IFTICFAPTNIL 336
DB 185 RVFVGFLFPWMLLSYRGILRAVGSVSTERQEKAKIKRLALSIAIVLVCFAPYHLL 244

QY 337 VIHANYYYHNTD----SLVFMYLIALCLSLNSCLDPPLYFVMSK 378
DB 245 LRSRAVYLGHWDGCFEERFVSVAHSSLAFTSLNCAVDPILYCLVNE 291

RESULT 25
I55450
G protein-coupled P2 receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
C:Accession: I55450
R:Chang, K.; Hanaoka, K.; Kumada, M.; Takuwa, Y.
J. Biol. Chem. 270, 26152-26158, 1995
A:Title: Molecular cloning and functional analysis of a novel P2 nucleotide receptor.
A:Reference number: I55450; MUID:96064682; PMID:7592819
A:Accession: I55450
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <RES>
A:Cross-references: GB:D63665; NID:g1066007; PIDN:BAA09816.1; PID:g1066008
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor

Query Match 15.6%; Score 333; DB 2; Length 328;
Best Local Similarity 27.9%; Pred. No. 5e-19;
Matches 87; Conservative 54; Mismatches 109; Indels 62; Gaps 12;

QY 104 RSSLSQVPIPIYILLFVVGVPNSIVTLWKL--SLRTKSTSLVIFHTNLAIAADLLFCVTL 161
DB 104 RSSLSQVPIPIYILLFVVGVPNSIVTLWKL--SLRTKSTSLVIFHTNLAIAADLLFCVTL 161

Db 21 REDFKLLPPVYVVLVGLPLNVCVIAQICASRRLTRS-AVYTLNLALADLLVACSL 79
 QY 162 PFKIAIYHLGNWVFGVWMCRIITTVFYGNMYCAIILTCMGINRYLATAHPPT-YOKLP 220
 Db 80 PLLIYNYARGDHPFGDLACRLVRFIFYANLHGSILFLICISFORVLGICHPLAPWHKRG 139
 QY 221 KRSFSLMCGIYVMVFLYMLPVILKQEVHLVHSEITTCCHVDVACESPS--SFRFY-Y 277
 Db 140 GRBAAVVCGVWVWVTAQCLPTAFAATG--IQNRNVCYDL-----SPFILSTRYLPY 192
 QY 278 FVSLAFPGELIPVLIIFCYTTLIHLKSKDRWLWG-----YIKAVLLILVIF 325
 Db 193 GMAITVIGFLPPTALLACYCRMARLQRD-----GPAGVQARRSKAARWAVVAVF 248
 QY 326 TICFAPNTIILVIHANYYYHNTDSLYF-----MYLIALCLGSLNSC 367
 Db 249 VISFLP-----PHIKTAYLAVRSTPGVSCPVIETFAAAYKGTFRFASANSV 295
 QY 368 LDPFL-YFVNSK 378
 Db 296 LDPILFYFTQOK 307

RESULT 26

IS1372
 angiotensin II receptor - turkey
 C:Species: Meleagris gallopavo (common turkey)
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Nov-1999
 C:Accession: I51372; P00449
 R:Murphy, T.J.; Nakamura, Y.; Takeuchi, K.; Alexander, R.W.
 Mol. Pharmacol. 44, 1-7, 1993
 A:Title: A cloned angiotensin receptor isoform from the turkey adrenal gland is pharmacologically identical to the human receptor
 A:Reference number: I51372; MUID:93341466; PMID:8341266
 A:Accession: I51372
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-359 <MUR>
 A:Cross-references: GB:I23203; NID:G349735; PIDN:AAA03560.1; PID:G349736
 R:Garsia, R.V.; McIlroy, P.J.; Kowaleki, K.I.; Tilly, J.L.
 Biochem. Biophys. Res. Commun. 191, 1073-1080, 1993
 A:Title: Isolation of turkey adrenocortical cell angiotensin II (AII) receptor partial cDNA
 A:Reference number: P00449; MUID:93221469; PMID:7916599
 A:Accession: P00449
 A:Molecule type: mRNA
 A:Residues: 53-91, 'C', 93-162, 'SSFTIVY', 171-291 <CAR>
 A:Cross-references: GB:S58041; NID:G299373; PIDN:AAB26041.1; PID:G299374
 A:Experimental source: liver
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: glycoprotein; phosphoprotein; transmembrane protein
 F:53-64/Domain: intracellular #status predicted <IN1>
 F:65-86/Domain: transmembrane #status predicted <TM1>
 F:87-102/Domain: extracellular #status predicted <EX1>
 F:103-124/Domain: transmembrane #status predicted <TM2>
 F:125-142/Domain: intracellular #status predicted <IN2>
 F:143-162/Domain: transmembrane #status predicted <TM3>
 F:163-192/Domain: extracellular #status predicted <EX2>
 F:193-214/Domain: transmembrane #status predicted <TM4>
 F:215-238/Domain: intracellular #status predicted <IN3>
 F:239-262/Domain: transmembrane #status predicted <TM5>
 F:263-274/Domain: extracellular #status predicted <EX3>
 F:275-291/Domain: transmembrane #status predicted <TM6>
 F:136/Binding site: phosphate (Ser) (covalent) #status predicted
 F:141,233/Binding site: phosphate (Thr) (covalent) #status predicted
 F:176,188/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.5%; Score 332; DB 2; Length 359;
 Best Local Similarity 26.9%; Pred. No. 6.6e-19;
 Matches 82; Conservative 65; Mismatches 140; Indels 18; Gaps 6;
 QY 87 EDSITSLHVNATIGYLRSLSTQVIPAIIYLLFVGVPSN--IVTLWLKLSLTKSLV 144
 Db 8 EETVKRIHDPVSG--RHSYIYIMVPTVYSIIIFIIGNLSLVIIYCYMKLKTVA-S 64

QY 145 IFTNLAIADLFCVTLFPKIAIYHLGNWVFGVWMCRIITTVFYGNMYCAIILTCMGI 204
 Db 65 IFLLNLALADLCLITLPLWAAATAMEYQWPGNCLKLASAGISFNLYASVFLTLCLSI 124
 QY 205 NRYLATAHPPTYOKLPKRSFSLMCGIYVMVFLYMLPVILKQEVHLVHSEITTCCHDV 264
 Db 125 DRYLAIVHPVKSIRRTMTFVARVTCIVWLGLAGVASLPVTHRNIFFAENLNMTCGFRY 184
 QY 265 DACESPSFRFYFVVSFLAFFGLIPFVIFCYTTLIHLKSKDRWLWG-----YIKAV 318
 Db 185 D--NNNTILRVCLGSLKSLGFLIPFLILTSYTLIWLTKKAYIQNRNTRDDIFKMI 242
 QY 319 LLILVIFTCFAPTNI-----ILVIHANYYYHNTDSLYFMYLIALCLGSLNSCLDPFLY 373
 Db 243 VAIVFFFSWIPHQVFTFLDLVLIQLHVITCKITDIVDTAMPFTICIAFYFNCLNPFFY 302
 QY 374 FVNSK 378
 Db 303 VFEGK 307

RESULT 27

JC2134
 angiotensin II receptor type 1A - rat
 N:Alternate names: AT1a receptor; AT1 receptor
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 24-Nov-1999
 C:Accession: JC2134; S15404; S20424; J01055
 R:Conchon, S.; Monnot, C.; Sirieix, M.E.; Bihoreau, C.; Corvol, P.; Clauser, E.
 Biochem. Biophys. Res. Commun. 199, 1347-1354, 1994
 A:Title: Synthetic cDNA encoding the rat AT1a receptor: a useful tool for structure-function studies
 A:Reference number: JC2134; MUID:94197726; PMID:8147879
 A:Accession: JC2134
 A:Molecule type: mRNA
 A:Residues: 1-359 <CON>
 A:Note: The amino acid sequence of this protein is not given
 R:Murphy, T.J.; Alexander, R.W.; Griendling, K.K.; Runge, M.S.; Bernstein, K.E.
 Nature 351, 233-236, 1991
 A:Title: Isolation of a cDNA encoding the vascular type-1 angiotensin II receptor.
 A:Reference number: S15404; MUID:91251901; PMID:2041570
 A:Accession: S15404
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-359 <MUR>
 A:Cross-references: GB:X62295; NID:G57773; PIDN:CAA44183.1; PID:G57774
 R:Iwai, N.; Inagami, T.
 FEBS Lett. 298, 257-260, 1992
 A:Title: Identification of two subtypes in the rat type I angiotensin II receptor.
 A:Reference number: S20423; MUID:92183879; PMID:1544458
 A:Accession: S20424
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-359 <IWA>
 R:Iwai, N.; Yamano, Y.; Chaki, S.; Konishi, F.; Bardhan, S.; Tibbette, C.; Sasaki, K.; I.
 Biochem. Biophys. Res. Commun. 177, 299-304, 1991
 A:Title: Rat angiotensin II receptor: cDNA sequence and regulation of the gene expression
 A:Reference number: JQ1055; MUID:91254291; PMID:2043116
 A:Accession: JQ1055
 A:Molecule type: mRNA
 A:Residues: 1-80, 'C', 82-108, 'T', 110-359 <IW2>
 A:Cross-references: GB:M74054; NID:G202918; PIDN:AAA40738.1; PID:G202919
 A:Experimental source: kidney
 C:Comment: AT1 receptor consists of two closely related AT1 isoforms of angiotensin II
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; glycoprotein; GTP binding; receptor; transmembrane protein
 F:27-86/Domain: transmembrane #status predicted <TM1>
 F:103-123/Domain: transmembrane #status predicted <TM3>
 F:147-167/Domain: transmembrane #status predicted <TM4>
 F:195-218/Domain: transmembrane #status predicted <TM5>
 F:240-263/Domain: transmembrane #status predicted <TM6>
 F:278-299/Domain: transmembrane #status predicted <TM7>
 F:4,176,188/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.3%; Score 327; DB 2; Length 359;
Best Local Similarity 28.0%; Pred. No. 1.6e-18;
Matches 90; Conservative 51; Mismatches 129; Indels 52; Gaps 8;

QY 87 EDSISLTHVNNATIGYLRSSLSQTQVIPAIYILLFVVGVSNN-----IVTLWKLRLTKSIS 142
DB 8 EDGKRIQDDCPKAG--RHSYIFWMIPTLSIIFVVGIFGNSLVVIVFYMKLKTVA-- 63

QY 143 LVIFHTNLAIADLLFCVTLTPFKIAYHLNGNNWVFGVMCRITTVVFGYNNYCAILLITCM 202
DB 64 -SVFLNLALADLCFLTLPLMAVYTAMEYRPFNGHLCCKIASVSFNLKAYEIQ 122

QY 203 GINRYLATAHPPTYQKLPKRSFSLMCGIWMVFLYMLPFVILKOEYHLVHSEITTCCHD 262
DB 123 SIDRYLAIVHPMKSLRLRTMLVAKVTCIIILWLAGLASLPAVHRNVYFIENITIVC-- 180

QY 263 VVDACESPSSFRYY-----FVSLAFPGFLIPFVIIIFCYTTLIHLKSKDRIW 311
DB 181 -----AFHYESRNSLTPIGLTKNLTGFLFPLIILTSYTLIWKALKKAYEIQ 229

QY 312 LG-----YIKAVLLILVITFCFAPTNIIL-----VIHANYYYHNTDSLYFMYL 356
DB 230 KNKPRNDIDFRIIMAILVLPFFFSWVPHQIFTLFDVLIQLGVHDCIKI-----SDIVDTAMP 285

QY 357 IALCLGSLNSCLDPFLYFVMSK 378
DB 286 ITICIAVFNCLNPLFYGLGK 307

RESULT 28
JH0621
A:Title: Cloning, characterization, and expression of two angiotensin receptor (AT-1) is
A:Reference number: JH0621; MUID:92287102; PMID:1599461
A:Accession: JH0621
A:Molecule type: DNA
A:Residues: 1-359 <SAS>
A:Cross-references: GB:S37484; NID:G249945; PIDN:AAB22269.1; PID:G249946
A:Experimental source: strain Balb/c
R:Yoshida, H.; Kakuchi, J.; Guo, D.F.; Furuta, H.; Iwai, N.; van der Meer-de Jong, R.; I
Biochem. Biophys. Res. Commun. 186, 1042-1049, 1992
A:Title: Analysis of the evolution of angiotensin II type 1 receptor gene in mammals (m
A:Reference number: JCI193; MUID:9235981; PMID:1497638
A:Accession: JCI193
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-6, 'I', 8-19, 'IS', 22-37, 'M', 39-133, 'K', 135-359 <YOS>
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F:37-57/Domain: transmembrane #status predicted <TM1>
F:65-85/Domain: transmembrane #status predicted <TM2>
F:103-123/Domain: transmembrane #status predicted <TM3>
F:145-165/Domain: transmembrane #status predicted <TM4>
F:201-220/Domain: transmembrane #status predicted <TM5>
F:241-261/Domain: transmembrane #status predicted <TM6>
F:286-306/Domain: transmembrane #status predicted <TM7>
F:4,176,188/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:331,338/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 15.3%; Score 326; DB 2; Length 359;
Best Local Similarity 28.0%; Pred. No. 2e-18;
Matches 90; Conservative 50; Mismatches 130; Indels 52; Gaps 8;

QY 87 EDSISLTHVNNATIGYLRSSLSQTQVIPAIYILLFVVGVSNN-----IVTLWKLRLTKSIS 142
DB 8 EDGKRIQDDCPKAG--RHSYIFWMIPTLSIIFVVGIFGNSLVVIVFYMKLKTVA-- 63

QY 143 LVIFHTNLAIADLLFCVTLTPFKIAYHLNGNNWVFGVMCRITTVVFGYNNYCAILLITCM 202

DB 64 -SVFLNLALADLCFLTLPLMAVYTAMEYRPFNGHLCCKIASVSFNLKAYEIQ 122

QY 203 GINRYLATAHPPTYQKLPKRSFSLMCGIWMVFLYMLPFVILKOEYHLVHSEITTCCHD 262

DB 123 SIDRYLAIVHPMKSLRLRTMLVAKVTCIIILWLAGLASLPAVHRNVYFIENITIVC-- 180

QY 263 VVDACESPSSFRYY-----FVSLAFPGFLIPFVIIIFCYTTLIHLKSKDRIW 311

DB 181 -----AFHYESRNSLTPIGLTKNLTGFLFPLIILTSYTLIWKALKKAYEIQ 229

QY 312 LG-----YIKAVLLILVITFCFAPTNIIL-----VIHANYYYHNTDSLYFMYL 356

DB 230 KNKPRNDIDFRIIMAILVLPFFFSWVPHQIFTLFDVLIQLGVHDCIKI-----ADIVDTAMP 285

QY 357 IALCLGSLNSCLDPFLYFVMSK 378

DB 286 ITICIAVFNCLNPLFYGLGK 307

RESULT 29
JH2492
G protein-coupled receptor 1 - rat
N:Alternate names: GPR-1
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 08-Oct-1999
C:Accession: JH2492
R:Marchese, A.; Cheng, R.; Lee, M.C.; Porter, C.A.; Heiber, M.; Goodman, M.; George, S.I
Biochem. Biophys. Res. Commun. 205, 1952-1958, 1994
A:Title: Mapping studies of two G protein-coupled receptor genes: An amino acid differer
A:Reference number: JH2492; MUID:95110347; PMID:7811287
A:Accession: JH2492
A:Molecule type: mRNA
A:Residues: 1-353 <MAR>
A:Cross-references: GB:S74702; NID:G786483; PIDN:AAB322978.1; PID:G786484
C:Superfamily: vertebrate rhodopsin
C:Keywords: glycoprotein; lipoprotein; phosphoprotein; receptor; thiolester bond; trans
F:74-94/Domain: transmembrane #status predicted <TM2>
F:112-133/Domain: transmembrane #status predicted <TM3>
F:134-135/Region: DR motif
F:155-175/Domain: transmembrane #status predicted <TM4>
F:209-229/Domain: transmembrane #status predicted <TM5>
F:246-266/Domain: transmembrane #status predicted <TM6>
F:295-306/Domain: transmembrane #status predicted <TM7>
F:14,273/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:150,231/Binding site: phosphate (Thr) (covalent) (by protein kinase A) #status predict
F:330/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 15.2%; Score 325; DB 2; Length 353;
Best Local Similarity 27.5%; Pred. No. 2.3e-18;
Matches 87; Conservative 59; Mismatches 110; Indels 60; Gaps 12;

QY 115 IYVLFVGVGVSNIIVTLWKLRLTKSISLIVFHTNLAIADLLFCVTLPPKIAHYLNGNNW 174

DB 44 LYALAFVLGIPGNAIVIFWFGFKKTKVTTLWFLNLAIADFLFVLPLIYISVALSPHW 103

QY 175 VFGVMCRITTVVFGYNNYCAILLITCMGINRYLATAHPPTYQKLPKRSF--SLLMCGIV 232

DB 104 PFGRLWCLKUNSLIAOLNMFSSVFFLTVISLDRYTHLIHPGLSH--PHRTLKNSLLVLFV 161

QY 233 WYVFLYMLPFVILK-----OYHLVHSEITTCCHDVVDACESPSSFRYY 277

DB 162 WLLASLLGGTPLYFRDTEVNNRIICNNFQYEL-----TLMRHV 203

QY 278 FVSLAP-FGFLIPFVIIIFCYTTLIHLK-----SKRIWGLYTKAVLLIIVITFCFA 330

DB 204 LTWVFLFGYLLPLLTMSSCYLCIFTKKQNILISSKHLWN-----ILSVVIAFMVCWT 258

QY 331 PTNII---LVTHANYYYHNTDSLYFMYLIAICLGSLNSCLDPFLYFVMSKVVDQLNFX 386

DB 259 PPHLSIWEISLHH--NSSFQNV--IQGGIPLSTGLAFLNSCLNPLIYIISKFF-QARPR 314

QY 387 SAMARPLXRRPRDIWE 402

Db 315 ASVAEVL---KRLWE 327
A:Accession: JH0622
A:Molecule type: DNA
A:Residues: 1-6, 'I', 8-165, 'H', 167-172, 'E', 174-204, 'V', 206-231, 'T', 233-238, 'F', 240-359 <S>
A:Cross-references: GB:S37491; NID:g249947; PIDN:AAB22270.1; PID:g249948
A:Experimental source: Balb/c
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein

RESULT 30
S26667
G protein-coupled receptor BLR1 - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C:Accession: S26667
R:Dobner, T.; Wolf, I.; Emrich, T.; Lipp, M.
Eur. J. Immunol. 22, 2795-2799, 1992
A:Title: Differentiation-specific expression of a novel G protein-coupled receptor from
A:Reference number: S26667; MUID:93049615; PMID:1425907
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-372 <DOB>
A:Cross-references: EMBL:X68149; NID:g29459; PIDN:CAA48252.1; PID:g29460
C:Genetics:
A:Gene: GDB:BLR1
A:Cross-references: GDB:136235; OMIM:601613
A:Map position: 15q26.1-15q26.1
A:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 15.2%; Score 324.5; DB 2; Length 372;
Best Local Similarity 28.6%; Pred. No. 2.7e-18;
Matches 90; Conservative 63; Mismatches 113; Indels 49; Gaps 14;

QY 87 EDSISTLHVNNTIGVLRSLSTQVTPAIYILLFVGVPSNIVTLWKLSLRTK---SISL 143
Db 29 DTSLVENHLCPATEGELMASFKAVFVPVAYSLIFLLGVGNVLVILERHRTSRSTET 88
QY 144 VIFHTNLAIADLLFCVTLPPKIAIYHLNGNNWVGEVMCRITTVFVGNMYCAILLTCMG 203
Db 89 FLRH--LAVADLLVILPFAVAGVG--WVLTGTLCKIVIAHLKVNFCYCSLLLACTA 144
QY 204 INRYLATAHPFTYOKLPKRSFSL-LMCGIWMVFLYMLP---FVILKOEYHLVHSEITT 259
Db 145 VDRYLAIVHA-VHAYRHRLLSHITCGITLWGLFLLAULPEILFAKVSQGH--NNSLPR 201
QY 260 C-HDVVDACESPSP--RFYFVSLAFPGFLIPFVILIIICVYTLIHLKSKOR--IWLGY 314
Db 202 CTFSENQAEATHAWFTSRFLYHA---GFLPLVMGWCYGVGVRHLRQAORRQORQA 257
QY 315 IKAVLLILVITFCFAPTNIILVIHANYHHNTDSLYFM-----YLIALC 360
Db 258 VRVAILVTSIFFLCWSPYHIVFL-----DTLRLKAVDNTCKLNGSLPVAITWC 307
QY 361 --LGSLNSCLDPLY 373
Db 308 EFLGLAHCLNPMLY 322

RESULT 31
JC1194
angiotensin II receptor 1B - mouse
C:Species: Mus musculus (house mouse)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Nov-1999
C:Accession: JC1194; JH0622
R:Yoshida, H.; Kakuchi, J.; Guo, D.F.; Furuta, H.; Iwai, N.; van der Meer-de Jong, R.; I
Biochem. Biophys. Res. Commun. 186, 1042-1049, 1992
A:Title: Analysis of the evolution of angiotensin II type 1 receptor gene in mammals (m
A:Reference number: JC1193; MUID:92359981; PMID:1497638
A:Accession: JC1194
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-359 <YOS>
R:Sasamura, H.; Hein, L.; Krieger, J.E.; Pratt, R.E.; Kobilka, B.K.; Dzau, V.J.
Biochem. Biophys. Res. Commun. 185, 253-259, 1992
A:Title: Cloning, characterization, and expression of two angiotensin receptor (AT-1) is
A:Reference number: JH0621; MUID:92287102; PMID:1599461

A:Accession: JH0622
A:Molecule type: DNA
A:Residues: 1-6, 'I', 8-165, 'H', 167-172, 'E', 174-204, 'V', 206-231, 'T', 233-238, 'F', 240-359 <S>
A:Cross-references: GB:S37491; NID:g249947; PIDN:AAB22270.1; PID:g249948
A:Experimental source: Balb/c
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein

F:30-53/Domain: transmembrane #status predicted <TM1>
F:65-90/Domain: transmembrane #status predicted <TM2>
F:103-123/Domain: transmembrane #status predicted <TM3>
F:147-166/Domain: transmembrane #status predicted <TM4>
F:195-218/Domain: transmembrane #status predicted <TM5>
F:240-263/Domain: transmembrane #status predicted <TM6>
F:278-300/Domain: transmembrane #status predicted <TM7>
F:4,176,188/Binding site: carboxylate (Asn) (covalent) #status predicted
F:141/Binding site: phosphate (Thr) (covalent) #status predicted
F:331,338,348/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 15.2%; Score 324; DB 2; Length 359;
Best Local Similarity 28.1%; Pred. No. 2.8e-18;
Matches 91; Conservative 52; Mismatches 125; Indels 56; Gaps 9;

QY 87 EDSISTLHVNNTIGVLRSLSTQVTPAIYILLFVGVPSN---IVTLWKLSLRTKSI 142
Db 8 EDSIKRIQDDCPKAG--RHNIFVIMPIYLSIFVGVFGNSLVIVIVFYMKLKTVA-- 63
QY 143 LVIFHTNLAIADLLFCVTLPPKIAIYHLNGNNWVGEVMCRITTVFVGNMYCAILLTCM 202
Db 64 -SVFLLNLALADLCFLLTLPLWAVTAMEYQWPGFNLCKIASASVSFNLYASVFLTLCL 122
QY 203 GINRYLATAHPFTYOKLPKRSFSLMCGIWMVFLYMLPFFVILKOEYHLVHSEITTCMD 262
Db 123 SIDRYLAIVHPKSLRRLTMLVAKVTCIIWLMAGLASLPVIRYRVFIANTNITVC-- 180
QY 263 VVDACESPSPRFY-----FVSLAFPGFLIPFVILIIICVYTLIHLKSKDRIW 311
Db 181 -----AFHESQNSTLPIGLTKNLIGFEPFVILITSTLTKWALKKAYKIQ 229
QY 312 -----LGYIKAVLLILVITFCFAPTNI-----ILVIHANYHHNTDSLYFM 354
Db 230 KNIPRNDIGRI--IMAVLPFFFSWPHQIFSLFDVLQIGVHDCI---ADVVDTA 283
QY 355 YLIALCLSLNSCLDPLYFVNSK 378
Db 284 MBTITICVFNCLNPLFYGLGK 307

RESULT 32
IS3033
G protein-coupled receptor - human
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: IS3033
R:Heiber, M.; Docherty, J.M.; Shah, G.; Nguyen, T.; Cheng, R.; Heng, H.H.Q.; Marchese, J
DNA Cell Biol. 14, 25-35, 1995
A:Title: Isolation of three novel human genes encoding G protein-coupled receptors.
A:Reference number: IS3033; MUID:95134353; PMID:7832990
A:Accession: IS3033
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-354 <RES>
A:Cross-references: GB:JL36148; NID:g598152; PIDN:AAA63180.1; PID:g598153
C:Superfamily: G protein-coupled receptor 4
C:Keywords: G protein-coupled receptor

Query Match 15.1%; Score 321.5; DB 2; Length 354;
Best Local Similarity 27.6%; Pred. No. 4.4e-18;
Matches 79; Conservative 59; Mismatches 113; Indels 35; Gaps 6;

QY 113 PAIYILLFVGVPSNIVTLWKLSLRTKSI-LVIFHTNLAIADLLFCVTLPPKIAIYHLNG 171
Db 21 PSLYIFVIGVGLPTNCLALWAAAYRQVQRNELGVIMLNLSIADLLYICTLPLWVDFLHH 80

Db 123 SDRYLAIIVHPMKSLRRRTMLVAKVTCIIIIWLAGLASLPAIIHRNVFFIENTNITVC-- 180
QY 263 VVDACESPSFRFY-----FVSLAFPGFLIPFVIIIFCVTTIIHLKSKDRITW 311
Db 181 -----AFHYESQNSTLPIGLGKTNILGFLFPFLIILTSYTLIWKALKKAYEIQ 229
QY 312 LG-----YIKAVLLILVITTCFAPTNIIL-----VIHANYYYHNTDSLYFMVL 356
Db 230 KNKPRNDDIFKIMAVLVLFFFSVPHQIFTELDVLQIQLGVHDCRI-----ADIVDTAMP 285
QY 357 IALCLGSLNSCLDPPFLYFVMSK 378
Db 286 ITICIAFYFNCLNPLFYGLGK 307

RESULT 36
S15403
angiotensin II receptor type 1 - bovine
C:Species: Bos primigenius indicus x Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Nov-1999
C:Accession: S15403
R:Sasaki, K.; Yamano, Y.; Bardhan, S.; Iwai, N.; Murray, J.J.; Hasegawa, M.; Matsuda, Y.
Nature 351, 230-233, 1991
A:Title: Cloning and expression of a complementary DNA encoding a bovine adrenal angiotensin II receptor type 1
A:Accession: S15403
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-359 <SAS>
A:Cross-references: GB:X62294; NID:943; PIDN:CAA44182.1; PID:944
C:Superfamily: vertebrate rhodopsin

Query Match 15.0%; Score 320; DB 2; Length 359;
Best Local Similarity 27.6%; Pred. No. 5.9e-18;
Matches 85; Conservative 50; Mismatches 103; Indels 70; Gaps 8;

QY 111 VIPAIYILFVVGVPNS-----IVTLWKLRLTKSISLVIFHTNLAIADLLFCVTLPEFKIA 166
Db 30 MIPTLYSIIFVVGIFGNSLWVIYFYMKLKTVA---SVFLNLALADLCFLTLPLMAV 86
QY 167 YHLGNMNVFGEVWCRTITVVFVGNMYCAIILITCMGINRYLATAHPFTYQKLPKESPSL 226
Db 87 YTAMEYRVPFNGVYLCIASASVSFNLYASVELTCLSDIDRYLAIVHPMKSLRRRTMLVAK 146
QY 227 LMGCIWVMVFLYVVGVPNS-----IVTLWKLRLTKSISLVIFHTNLAIADLLFCVTLPEFKIA 166
Db 147 VTICIIWLAGLASLPTIIHRNVFFIENTNITVC-----AFHYESQNSTLVP 193
QY 278 --FVSLAFPGFLIPFVIIIFCVTTIIHLKLS-----KDIRIWLGVYKAVLLILVIF 325
Db 194 GLGLTKNIGLFLPFLIILTSYTLIWKALKKAYEIQKNPKRDIIF---KIILAIVLFF 249
QY 326 TICFAPTNIILVIHANYYYHNTDSLYFMVL-----IALCLGSLNSCLDP 370
Db 250 FFSVPHQIFTEF-----DVLIIQLGLRDCKIEDIVDTAMPITICLAYFNNSCLNP 299
QY 371 FLYFVMSK 378
Db 300 LFYGLGK 307

RESULT 37
JC5653
G protein-coupled receptor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 17-Nov-1997 #sequence_revision 17-Nov-1997 #text_change 21-Jul-2000
C:Accession: JC5653
R:Ferrer-Martinez, A.; Felipe, A.; Mata, J.F.; Casado, F.J.; Pastor-Anglada, M.
Biochem. Biophys. Res. Commun. 238, 107-112, 1997
A:Title: Molecular cloning of a bovine renal G-protein coupled receptor gene (BRGR): Res
A:Reference number: JC5653; MUID:97445134; PMID:9299461
A:Accession: JC5653
A:Molecule type: mRNA

A:Residues: 1-361 <FER>
A:Cross-references: GB:U88366; NID:g2827875; PIDN:AA05611.1; PID:g2827876
A:Experimental source: renal epithelial cell
C:Superfamily: G protein-coupled receptor 4
C:Keywords: glycoprotein; phosphoprotein
F:24-46/Domain: transmembrane #status predicted <TM1>
F:58-79/Domain: transmembrane #status predicted <TM2>
F:96-117/Domain: transmembrane #status predicted <TM3>
F:137-153/Domain: transmembrane #status predicted <TM4>
F:189-207/Domain: transmembrane #status predicted <TM5>
F:229-253/Domain: transmembrane #status predicted <TM6>
F:265-289/Domain: transmembrane #status predicted <TM7>
F:3,8,62/Binding site: carbonylrate (Asn) (covalent) #status predicted
F:134,218,293,313/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status pre
F:204,221,328/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pre

Query Match 15.0%; Score 319.5; DB 2; Length 361;
Best Local Similarity 26.7%; Pred. No. 6.5e-18;
Matches 76; Conservative 65; Mismatches 107; Indels 37; Gaps 8;

QY 113 PAIYILLFVVGVPNSIVTLWKLRLTKSIS-LVIFHTNLAIADLLFCVTLPEFKIAHYLNG 171
Db 25 PVVYVTVLVVGFPANCLSLYCYLQIKARNELGVVLCNLTVADLFYICSLPFWLQVLOH 84
QY 172 NNWVGEVWCRTITVVFVGNMYCAIILITCMGINRYLATAHPFTYQKLPKRSFSLMCGI 231
Db 85 DWSHDDLSCQVCGILLYENIYISVGLCCISIDRYLAVAHFPHQFRTLKAAMGVSA 144
QY 232 VVMVFLYMLPFVILKQBYHLVHSEITTCHEVDVAC-----ESPSSFRFYFYVSLAFF 284
Db 145 IWKELLTSI-----YFLMHSEVVEDADRHRVCFEHPLEPRQGINYRFL-----V 192
QY 285 GFLIPFVIIIFCYTTLIHLK-----SKDIRIWLGVYKAVLLILVITTCFAPTNIIL 336
Db 193 GFLPFCILLASRYGILRAVRSHCTQSKRDKIQO----RLVLSVIVFLACFLPYHVL 248
QY 337 VIHANYYYHNTD---SLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 249 LVR--SLWESSCDFAKGIFNAYHFSLTLTTFNCVADPVLVCFVSE 291

RESULT 38
S68208
G protein-coupled receptor 12A - human
C:Species: Homo sapiens (man)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 29-Sep-1999
C:Accession: S68208
R:An, S.; Tsai, C.; Goetzl, E.J.
FEBS Lett. 375, 121-124, 1995
A:Title: Cloning, sequencing and tissue distribution of two related G protein-coupled re
A:Reference number: S68207; MUID:96087098; PMID:7498459
A:Accession: S68208
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-365 <ANS>
A:Cross-references: EMBL:U35398; NID:gi015418; PIDN:AAA79060.1; PID:gi015419
C:Superfamily: G protein-coupled receptor 4
C:Keywords: G protein-coupled receptor

Query Match 14.9%; Score 318.5; DB 2; Length 365;
Best Local Similarity 26.8%; Pred. No. 7.9e-18;
Matches 77; Conservative 67; Mismatches 102; Indels 41; Gaps 10;

QY 113 PAIYILLFVVGVPNSIVTLWKLRLTKSIS-LVIFHTNLAIADLLFCVTLPEFKIAHYLNG 171
Db 25 PVVYVTVLVVGFPANCLSLYCYLQIKARNELGVVLCNLTVADLFYICSLPFWLQVLOH 84
QY 172 NNWVGEVWCRTITVVFVGNMYCAIILITCMGINRYLATAHPFTYQKLPKRSFSLMCGI 231
Db 85 DWSHDDLSCQVCGILLYENIYISVGLCCISIDRYLAVAHFPHQFRTLKAAMGVSA 144
QY 232 VVMVFLYMLPFVILKQBYHLVHSEITTCHEVDVAC-----TTC--HDVVDAACESPSFRFYVSLA 282

Db 145 IWAKELLTSI-----YFLMHEEVIEDENQHRVCFEHPYIQAWO--RAINYYRFL---- 191
QY 283 FFGFLPFPVLIIFCYTTLIHKL-----SKDRIWLGVIKAVILLIIVFTICAPTNI 334
Db 192 -VGFLPFCILLASYOGILRAVRSHGTSKRKQIQQ---RLVSTVVIIFLACFLPHV 246
QY 335 ILVIHHYYHNTD---SLYFMYLIJALCIGLSNCLDPLFPYFVMSK 378
Db 247 LLLVR--SVWEASCDPAKGVFNAYHFSLLLTSCNVADPVLICFVSE 291

RESULT 39
A53752
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: A53752
R;Prado, G.N.; Thomas, K.M.; Suzuki, H.; LaRosa, G.J.; Wilkinson, N.; Folco, E.; Navarro J. Biol. Chem. 269, 12391-12394, 1994
A;Title: Molecular characterization of a novel rabbit interleukin-8 receptor isotype.
A;Reference number: A53752; MUID:94230294; PMID:8175642
A;Accession: A53752
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-358 <PRA>
A;Cross-references: GB:L24445; NID:g437661; PIDN:AAA31378.1; PID:g437662
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 14.9%; Score 318; DB 2; Length 358;
Best Local Similarity 26.7%; Pred. No. 8.4e-18;
Matches 92; Conservative 74; Mismatches 116; Indels 62; Gaps 15;

QY 62 TFEPLSLDIEGTGAT---TTI-KACPDGSIISLHNNATIGVLRSSLSLSTQVIPAII 116
Db 10 SYEDF-FGDFSNYSYSTDPPTLLDSAPCRSESLT-----NSYVLLITY 53
QY 117 ILLFVGVPSN-IVTLWKLRLTKSISLVIFHNLAIDLFCVTLPEKIAVHLNGNNV 175
Db 54 ILVFLSLGNSLVMLVILYSRSTCVTDVYLNLAIDLFAITLPIWAASKVHG--WT 111
QY 176 FGEVMCRIITVWFYGNMYCAILITCMGINRYLATAHPTTYQKLPKRSFLMCGIVWM 235
Db 112 FGTPCKVWSLVKEVNFYSGILLACISVDRLAIVHA-TRTMIQKRLVKEICLSMVG 170
QY 236 VFVLMPLFVILKQYHLVHSEITTHDQVDVACESPSPSPFPYFVSLAFGLPFIPIIF 295
Db 171 SLILSILPILFLRNAIFPPNSS-PVCYE--DMGNSTAKRMVLRILPQTGFILPLVMLF 227
QY 296 CYTTLIHKLSKDRILWLG---IKAVLLIIVFTICFAPTNIILVIHHANYHHNTDSL 351
Db 228 CY---VFTLTLFQAHGQKRAMRVIFAVLLFLCWLFPYNLVLD-----TDIL 274
QY 352 YFMYLI-----AL-----CLGSLNSCLDPELY-FVMSK 378
Db 275 MRTHVIQETCERNRNDIDRALDATEILGLHSLCLNPIIYAFIGQK 318

RESULT 40
JC4800
P2Y6 receptor - human
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 17-Nov-2000
C;Accession: JC4800; G02514
R;Communi, D.; Parmentier, M.; Boeynaems, J.M.
Biochem. Biophys. Res. Commun. 222, 303-308, 1996
A;Title: Cloning, functional expression and tissue distribution of the human P2Y6 receptor
A;Reference number: JC4800; MUID:96222498; PMID:8670200
A;Accession: JC4800
A;Molecule type: mRNA
A;Residues: 1-328 <COM>
A;Cross-references: EMBL:X97058
A;Experimental source: placenta

R;Hammet, F.; Southey, M.C.; Somers, G.R.; Hutchins, A.M.; Venter, D.J.
submitted to the EMBL Data Library, March 1996
A;Reference number: H01373
A;Accession: G02514
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 'M', 4-328 <HAM>
A;Cross-references: EMBL:U52464; NID:g1407632; PIDN:AAB03572.1; PID:g1407633
C;Genetics:
A;Gene: P2Y6
C;Superfamily: ATP receptor P2u
C;Keywords: glycoprotein; placenta; receptor; transmembrane protein
F;26-52/Domain: transmembrane #status predicted <TM1>
F;63-86/Domain: transmembrane #status predicted <TM2>
F;104-122/Domain: transmembrane #status predicted <TM3>
F;143-167/Domain: transmembrane #status predicted <TM4>
F;193-216/Domain: transmembrane #status predicted <TM5>
F;241-264/Domain: transmembrane #status predicted <TM6>
F;283-305/Domain: transmembrane #status predicted <TM7>
F;5,173/Binding site: carbonyl (Asn) (covalent) #status predicted

Query Match 14.9%; Score 317.5; DB 2; Length 328;
Best Local Similarity 27.7%; Pred. No. 8.4e-18;
Matches 83; Conservative 53; Mismatches 125; Indels 39; Gaps 9;

QY 104 RSSLSLSTQVIPAIIYILLFVGVPSNIVTLWKLRLTKSISLVIFHT-NLAIALDLFCVTL 162
Db 21 RENFKQLLLPPVYSAVLAAGPLNICVITQCTSRALTRTAVYTLNALADLLYACSLP 80
QY 163 FKIAVHLNGNNVFEVGMCRITTVVYGNMYCAILITCMGINRYLATAHPTT-YOKLPK 221
Db 81 LLIYNYAQGDHWPFGDFACRLVRLFYANLHGSILFLTCISFORVILGICHPLAPWHKRG 140
QY 222 RSFSLMCGIIVWMVFLYMLFPVILKQYHLVHSEITTHDQVDVACESPSPSPFPYF--- 278
Db 141 RRAAWLVCVAVMLAVTTQCLTPAATG--IQNRNVTVCYDL-----SPPALATHYMPYG 193
QY 279 VSLAFPGFLIPFVLIIFCYTTLIHKLSKDRILWLGVIKAV-----LLILVIFT 326
Db 194 MALTVIGFLFPAALLACYCLLACLQD---GPAEPVAQERKGAARMVAVVAAFA 249
QY 327 ICFAPTNIILVIHHANYHHNTDSLIFMYLIJALC-----LGSLSNCLDPLFPYFVMSK 378
Db 250 ISFLPFHITKTAYLA---VRSTPGVPCVLEAFAAAYKGRPFASANSVLDPIILFYFTQK 306

Search completed: June 24, 2003, 12:06:34
Job time : 28.9681 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:05:46 ; Search time 32.4601 Seconds
(without alignments)
1356.747 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLTYXHPVAGSQDKMKIL.....AMARPLXRRPRDIWIDHAW 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues
Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US05_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	509	23.8	425 10	US-09-782-980-80	Sequence 80, Appli
2	509	23.8	425 10	US-09-884-430-4	Sequence 4, Appli
3	493	23.1	359 9	US-10-094-417-10	Sequence 10, Appli
4	493	23.1	359 9	US-09-782-974C-76	Sequence 76, Appli
5	493	23.1	359 10	US-09-739-151-2	Sequence 2, Appli
6	489	22.9	359 9	US-10-190-469-1	Sequence 1, Appli
7	444.5	20.8	359 10	US-09-943-718-6	Sequence 6, Appli
8	431.5	20.2	361 9	US-10-222-024-2	Sequence 2, Appli
9	431.5	20.2	361 9	US-10-251-385-78	Sequence 78, Appli
10	428.5	20.1	361 9	US-10-251-385-206	Sequence 206, App
11	426.5	20.0	348 10	US-09-827-937A-17	Sequence 17, Appli
12	399	18.7	362 9	US-10-092-135-3	Sequence 3, Appli
13	399	18.7	362 9	US-09-779-679-28	Sequence 28, Appli
14	398.5	18.7	367 9	US-09-828-478-6	Sequence 6, Appli
15	395	18.5	339 9	US-09-828-478-4	Sequence 4, Appli
16	395	18.5	339 9	US-10-251-385-32	Sequence 32, Appli
17	395	18.5	339 10	US-09-848-889-12	Sequence 12, Appli
18	395	18.5	339 10	US-09-788-133-2	Sequence 2, Appli
19	395	18.5	362 9	US-10-092-135-4	Sequence 4, Appli

20	393	18.4	537	9	US-10-311-956-4	Sequence 4, Appli
21	389	18.2	339	9	US-10-251-385-182	Sequence 182, App
22	388	18.2	366	9	US-09-779-679-25	Sequence 25, Appli
23	378.5	17.7	330	10	US-09-826-791-2	Sequence 2, Appli
24	378.5	17.7	346	9	US-09-828-478-2	Sequence 2, Appli
25	378.5	17.7	346	9	US-09-779-679-2	Sequence 2, Appli
26	378.5	17.7	346	9	US-09-779-679-26	Sequence 26, Appli
27	378.5	17.7	346	10	US-09-826-791-6	Sequence 6, Appli
28	378.5	17.7	346	10	US-09-866-230-7	Sequence 7, Appli
29	378.5	17.7	346	10	US-09-866-230-9	Sequence 9, Appli
30	378	17.7	337	9	US-09-828-478-5	Sequence 5, Appli
31	378	17.7	337	10	US-09-866-230-8	Sequence 8, Appli
32	377.5	17.7	373	9	US-10-092-135-7	Sequence 7, Appli
33	377	17.6	302	9	US-10-024-494-30	Sequence 30, Appli
34	377	17.6	337	9	US-09-779-679-27	Sequence 27, Appli
35	376.5	17.6	373	9	US-10-092-135-6	Sequence 6, Appli
36	372	17.4	373	9	US-10-092-135-5	Sequence 5, Appli
37	371.5	17.4	360	10	US-09-943-718-4	Sequence 4, Appli
38	370	17.3	337	9	US-10-167-192-3	Sequence 3, Appli
39	369.5	17.3	337	9	US-10-023-775B-2	Sequence 2, Appli
40	369.5	17.3	337	9	US-10-270-144-2	Sequence 2, Appli
41	369.5	17.3	337	9	US-10-188-405-8	Sequence 8, Appli
42	369.5	17.3	337	9	US-09-885-453-1	Sequence 1, Appli
43	369.5	17.3	337	9	US-10-079-384-14	Sequence 14, Appli
44	369.5	17.3	337	10	US-09-943-798-4	Sequence 4, Appli
45	369.5	17.3	341	9	US-10-270-587-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-782-980-80
; Sequence 80, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran M.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, AND
; TITLE OF INVENTION: STWST PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-121CP
; CURRENT APPLICATION NUMBER: US/09/782,980
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/02125
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/448,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: 60/117,580
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 09/014,195
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/014,348
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/086,892
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/296,208
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 09/063,950
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 09/561,381
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,810
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/087,121

```
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/672,721
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/049,799
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-380-80

Query Match      23.8%; Score 509; DB 10; Length 425;
Best Local Similarity 31.2%; Pred. No. 2e-36;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

QY 20 LILVAAGLLFLPVTVCOSGINVSDNSAKP-----TLTIKSF-NGGPONTPEEPFLSDI 71
DB 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDRSFLLRPNPNKYEPPWEDEE 60
QY 72 EGWTGAT-----TTIKACPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNS 127
DB 61 KNESGLTEYRLVSNKSPLOKQLPAPISEDASGYLTSSWLTFLFVPSVYTGCVFVSLPLN 120
QY 128 I--VTLWKLRLTKSISLIVFHTNLAIADLLFCVTLFPFKIAYHLNGNNVFGVEMCRITT 185
DB 121 IMAIVFILMKVKK--PAVVYMLHLATADVLVSVLPFKISYFSGSDWQFSGELCRFVT 179
QY 186 VVFYGNMYCAIILTCMGINRYLATAHP---FTYQKLPKRSFSLLMCGIWMVVMFLYMLP 242
DB 180 AAFYCNMYASILLMTVISIDRFVAVVPMQSLSWRTLGRASFT---CLAIWALAIAGVVP 236
QY 243 FVILKQEHVHLVHSEITTCCHVDVVDACESPSFRFYFVSLAFFGLIPFVIIIFCYTTLIH 302
DB 237 LVLKEQTIQVGLNITTCCHVDVNLTEGYYAYYFSAFSAVF--FFVPLLIISTVCYVSIIR 295
QY 303 KLKS-----KDRIWLGVIKAVLL---ILVIFTICFAPTNIILVIHANYYYH--NTD 349
DB 296 CLSSAVANRSKSR-----ALFLSAVFCIFICGPTNVLIIAHY-SFLSHTSTTE 347
QY 350 SLYFMYLIALLCLGSLNSCLDPFLYFVMS 377
DB 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 3
US-10-094-417-10
; Sequence 10, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiagang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: NO. US20030045685A1el Receptors
; FILE REFERENCE: 018781-008110US
; CURRENT APPLICATION NUMBER: US/10/094,417
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/802,803
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/276,649
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR341
US-10-094-417-10

Query Match      23.1%; Score 493; DB 9; Length 359;
Best Local Similarity 35.7%; Pred. No. 4.1e-35;
Matches 106; Conservative 53; Mismatches 116; Indels 22; Gaps 6;

QY 96 NNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSITVTLWKLRLTKSIS-LVIFHTNLAIAD 154
DB 10 DNATLQMLRNPAFAVALPVVYSLVAASIPGNLFSLWVLCRRMGPRSPSVIFMINLSVTD 69
QY 155 LFLCVTLFPFKIAYHLNGNNVFGVEMCRITTTFVYGNMYCAIILTCMGINRYLATAHPF 214
DB 70 LMLASVLPFQIYYHCNRHHWFGVLLCNVTVVAFYANMYSSILLTMTICISVERELGVLYPL 129
QY 215 TYQKLPKRSFSLLMCGIWMVVMFLYMLPFLVILKQEHVHLVHSEITTCCHVDVVDACESPS 271
DB 130 SSKRRRRRYAAACAGTWTLLLTALSPARTDLTYPVHALGIITCFDVLKWTMLPSVAM 189
```

Qy	272	--SFRYYVSVLAFGFLIPWIIIFCVTTLIHL-----KSDRIWLGIKAVLLI	321
Db	190	WAVLFTTIFIL-----FLIPFVTVACVATYATLKLRTTEAHGREQRRAVGLAAVVLL-	244
Qy	322	LVIFTCFAPNTIILVIHANYYYHNTDSLFYMYLIALCLGLNSCLDPLPYFVMSK	378
Db	245	--AFTVCFAPNNFVLLAHIVSRLLFG--KSYVHVYKLTCLICLNNCLDPLPYVFASR	298

```

RESULT 4
US-09-782-974C-76
: Sequence 76, Application US/09782974C
: Publication No. US20030082534A1
: GENERAL INFORMATION:
: APPLICANT: Vogeli, Gabriel
: APPLICANT: Lind, Peter
: APPLICANT: Wood, Linda S.
: APPLICANT: Parodi, Luis A.
: TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
: FILE REFERENCE: 411USPHRM311
: CURRENT APPLICATION NUMBER: US/09/782, 974C
: CURRENT FILING DATE: 2002-09-04
: PRIOR APPLICATION NUMBER: 60/165, 838
: PRIOR FILING DATE: 1999-11-16
: PRIOR APPLICATION NUMBER: 09/714, 449
: PRIOR FILING DATE: 2000-11-16
: PRIOR APPLICATION NUMBER: 60/198, 568
: PRIOR FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: 60/166, 071
: PRIOR FILING DATE: 1999-11-17
: PRIOR APPLICATION NUMBER: 60/166, 678
: PRIOR FILING DATE: 1999-11-19
: PRIOR APPLICATION NUMBER: 60/173, 396
: PRIOR FILING DATE: 1999-12-28
: PRIOR APPLICATION NUMBER: 60/184, 129
: PRIOR FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: 60/185, 421
: PRIOR FILING DATE: 2000-02-28
: PRIOR APPLICATION NUMBER: 60/185, 554
: PRIOR FILING DATE: 2000-02-28
: PRIOR APPLICATION NUMBER: 60/186, 530
: PRIOR FILING DATE: 2000-03-02
: Remaining Prior Application data removed - See File Wrapper or PALM..
: NUMBER OF SEQ ID NOS: 192
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 76
: LENGTH: 359
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-782-974C-76

```

Query Match	23.1%;	Score 493;	DB 9;	Length 359;
Best Local Similarity	35.7%;	Pred. NO. 4.le-35;		
Matches	106;	Conservative 53;	Mismatches 116;	Indels 22; Gaps 6
Qy	96	NNATIGYLRSSLSSTGVIPAIYLLFVGVGPSNIVTLKLSLRTKTSIS-LVIFHTNIIAIAID	154	
Db	10	DNATLQMLRNPAIAVALPVVYLSLVAASVPGNLSLVLWVLCRMCGRPSVSVFMINLSVTD	69	
Qy	155	LLECVTLPPKIAHYLGNNGNWGEVWCRRITTVVFGYNNMYCAILLTCMGINRYLATAHPF	214	
Db	70	LMLASVLPFQIYVHCNRRHHVFGVLLCNVTVAFVANNMYSSILMTCTGISVERELGLVLYPL	129	
Qy	215	TYOKLPRKRSLLMCGIVVMVFLYMLPFVILKQEYHLVHSEITTHCDVVDVADACESPS---	271	
Db	130	SSKWRERRRYAVAACAGTWLLLLLTALSPARTDLTPYHALGIITCFDVLKWTMLPSVAM	189	
Qy	272	--SFRYYFVSVLAFFGLPIPFVIIIFCVTTLIHK-----KSKDIRIWLGIYKAVLLI	321	
Db	190	WAYFLEFTIIFL-----FLIPFVITVACVYATAILKLLRTEEAHGREQRRRAVGLAAVVLL-	244	
Qy	322	LVFTTICFAPTTNIIILVIHHANNYYHYNTDSLYFMYLIALCLGSLMSCIDPPIFYFVMSK	378	

```

Db      245  --AFVTCFAPNNFVLLAHIVSRFLFYG-KSYHYHVYKLTLCUSCLNCLDPPFYVPASR 298
RESULT 5
US-09-739-151-2
; Sequence 2, Application US/09739151
; Patent No. US20010029032A1
; GENERAL INFORMATION:
; APPLICANT: Yuan Zhu
; APPLICANT: Xiaotong Li
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: PAUL, A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: POLYPEPTIDE
; FILE REFERENCE: GP-70567-C1
; CURRENT APPLICATION NUMBER: US/09/739,151
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/413,534
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/103,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 359
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-739-151-2

Query Match      23.1%; Score 493; DB 10; Length 359;
Best Local Similarity 35.7%; Pred. No. 4.1e-35;
Matches 106; Conservative 53; Mismatches 116; Indels 22; Gaps

Qy      96  NNATIGYLRSSLTGVIPAIYILLFVVGVPSPNIVTLWKLSLRTKSIS-LVIFHTNLAIAD
          :|||: |||: :|||: |||: |||: |||: |||: |||: |||: |||:
Db      10  DNATLQMLRNPAIAVALPWYSLVAASVPGNLSFLWLCHRRMGPRSPSVIFMNLNSVTD

Qy      155  LFLCVTLPPKIAHYLNGNNWTFGEVWCRIITTVFYGNMYCAILLTTCWNGINRYLYATAHPF
          :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db      70  LMLASVLPFQIYHYCHNRHHWFGVLLCNVTVAFYANNMYSSILMTTCISVERFLGVLYLP

Qy      215  TVQKLPKRSFSLMCGIVVMVFLYMLPFVILKQBYHLVHSEIITTHDVVVDACESSP-----
          :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db      130  SKRWRRRRYAAACAGTWLLLTALSPARTDITVPVHALGIITCFDVLKWTWMLPSVAM

Qy      272  --SFREYFFVSLAFFGFLPIFVIIFFCYTTLLIHL-----KSKDIRMLGWIKYIKAVLLI
          :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db      190  WAVFLFTIFILL---FLIPVITVACYTATILKLLRTEEAHGQRRRAVGLAAVULL-

Qy      322  LVITFTCFAPTNILIVIHIANHYHYHNTDSLTFMYLIALCLGSLNSCLDPLFYFVMSK 378
          :|||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db      245  --AFVTCFAPNNFVLLAHIVSRFLFYG-KSYHYHVYKLTLCUSCLNCLDPPFYVPASR 298

```

```

RESULT 6
US-10-190-469-1
; Sequence 1, Application US/10190469
; Publication No. US20030013155A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 14400 Receptor, A No. US20030013155A1 G-Protein Co
; FILE REFERENCE: 5800-7, 035800/169196
; CURRENT APPLICATION NUMBER: US/10/190,469
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: US/09/137,063A'
; PRIOR FILING DATE: 1999-08-20
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-190-469-1

```

[illegible]

```

1  RESULT 7
2  US-09-943-718-6
3  ; Sequence 6, Application US/09943718
4  ; Patent No. US20020103361A1
5  ; GENERAL INFORMATION:
6  APPLICANT: Huffine, Constance F.
7  Rossi, Devora L.
8  Capone, Myriam
9  Hedrick, Joseph A.
10 Vicari, Alain
11 Gorman, Daniel M.
12 Zlotnik, Albert
13
14 TITLE OF INVENTION: Mammalian Chemokines; Receptors;
15
16 NUMBER OF SEQUENCES: 8
17 CORRESPONDENCE ADDRESS:
18 ADDRESSEE: DNAX Research Institute
19 STREET: 901 California Avenue
20 CITY: Palo Alto
21 STATE: California
22 COUNTRY: USA
23 ZIP: 94304-1104
24
25 COMPUTER READABLE FORM:
26 MEDIUM TYPE: Floppy disk
27 COMPUTER: IBM PC compatible
28 OPERATING SYSTEM: PC-DOS/MS-DOS
29 SOFTWARE: PatentIn Release #1.0, Version #1.30
30
31 CURRENT APPLICATION DATA:
32 APPLICATION NUMBER: US/09/943,718
33 FILING DATE: 30-Aug-2001
34 CLASSIFICATION: <Unknown>
35
36 PRIORITY INFORMATION DATA:
37 APPLICATION NUMBER: 09/009,817
38 FILING DATE: 20-APR-1998
39
40 ATTORNEY/AGENT INFORMATION:
41 NAME: Ching, Edwin P.
42 REGISTRATION NUMBER: 34,090
43 REFERENCE/DOCKET NUMBER: DX0588K
44
45 TELECOMMUNICATION INFORMATION:
46 TELEPHONE: (650)852-9196
47 TELEFAX: (650)496-1200
48
49 INFORMATION FOR SEQ ID NO: 6:
50
51 SEQUENCE CHARACTERISTICS:
52 LENGTH: 359 amino acids
53 TYPE: amino acid
54 TOPOLOGY: linear
55 MOLECULE TYPE: protein

```

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-943-718-6

Query Match      20.8%; Score 444.5; DB 10; Length 359;
Best Local Similarity 33.3%; Pred. No. 7e-31;
Matches 93; Conservative 57; Mismatches 110; Indels 19; Gaps 4

QY   108 STQVIPAIYILLFVGVSPSNIVTLWKLSLRKTSISLVIFHTNLAIAIDLFCFTVPFKIAY 167
    . : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   51 SPQLVPALYGLVAVGLPANGALWLATVRPRLPSTILLMNLAVADLLALLVLPPRLAY 110

QY   168 HLGNNWVFGEVWCRIITVVPGYNMYCAIILITCMGINRXLATAPETYOKLPKRSFSL 227
    . : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   111 HLRGQRPFPEACRVATAALYGHMGYSVILLAAVSUDRYLALVHLRARALRGORLTG 170

QY   228 MCGIVVMWFELYMLPFVILKQEYHVLHSEITTCHDVVDAC-----ESPSSFRFYF 278
    . : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   171 LCLVAMLSAATLALPLTHRQNRELL-APIACC--VMRCPWLSRTPTGERPSSAWLSWA 227

QY   279 VSLAFPGFLIPFVIIIFCYTTLIHKLKSORIMGYTKAVLLILVITICPAPNIILVI 338
    . : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   228 ASL-----PLLAMGICYCTTLRALAANGORYSHARLRTLALVLFSAVASFTPSNVLLVL 280

QY   339 HHANYYHYNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
    . : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   281 HYNPSPDPAWGNYGAYVPSLALSTLNSCVDPPFIYYVVS 319

RESULT 8
US-10-222-024-2
; Sequence 2, Application US/10222024
; Publication No. US20030104487A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Ltd. (GB)
; TITLE OF INVENTION: Neuropeptide receptor and uses thereof
; FILE REFERENCE: PCS22032
; CURRENT APPLICATION NUMBER: US/10/222,024
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: GB 0119920.7
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-024-2
```

Query Match	20.2%	Score	431.5	DB	9	Length	361		
Best Local Similarity	31.6%	Pred. No.	9.5e-30						
Matches	95	Conservative	64	Mismatches	107	Indels	35	Gaps	9
<hr/>									
Qy	102	YLRSSLSTQVIPAIIYLLFVVUGPSNIVTLWKLSLRTKSI-SLVIFHTNLAIADLLFCVT	160						
Db	24	YAHSTARIYMPHLSVLFVIGLVGNLLALVIVQNRKKINSTLYSLNVLISDILFTTA	83						
Qy	161	LPFKIAYHLGNWNVFEVGMCRITTVFVYGNMYCAILITCMGNRNYLATAHPFTYQKLP	220						
Db	84	LPTRIAIYANGFDWRIGDALCRITALVFYINTYAGVNFMTCLSDRFAIVVHPHRYNKIK	143						
Qy	221	KRGFSLLMCGIVVMVFLYMLPFI-----LKQEYHLVHSEITTCDDVVDACESPSPREFFY	277						
Db	144	RIEHAKGVCIFVWLIVFAQTLPLLINPMKQE-----AERITCMEYPNFEETKS--LPWI	196						
Qy	278	FVSIUAFGFLIPFVIIIFCYTTLIHK-----LKSORIMILGYIKAVLLILVIFPIC	328						
Db	197	LLGACFIGYVPLIIIIICYSQICCKLFRTAKQNPLTEKSGWNKKALNTIILIIIVFVFLC	256						
Qy	329	FAPNIIILVH-----HANY-----YHNTD-SLYFMVLIALLGLSUNSCIDPELYFVMS	377						
Db	257	FTPYHVIAIIQHMIKKLRFSNPLECSQRHSFQISLHF-----TVCULNFNCMDPFYIFPAC	312						

Qy 378 K 378
Db 313 K 313

RESULT 9

US-10-251-385-78
; Sequence 78, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-78

Query Match 20.2%; Score 431.5; DB 9; Length 361;
Best Local Similarity 31.6%; Pred. No. 9.5e-30;
Matches 95; Conservative 64; Mismatches 107; Indels 35; Gaps 9;

Qy 102 YLRSSLSTQVIPAIIYLLFVVGVPSNIVTLWKLRLTKSI-SLVIFHTNLAIADLLFCVT 160
Db 24 YAHSTARIWPLHYSLVFIIGLVGNLLALVIVQNRKKINSTLYSTNLVSDILFTTA 83
Qy 161 LPFKIAYHLNGNNWVGEVNCRIITVVFYGNMYCAIILTCMGINRYLATAHPTFYOKLP 220
Db 84 LPTRIAIYAMGFDWRIGDLCRITALVFYINTYAGVNFMTCLSIDRFAIVVHPLRYNKK 143
Qy 221 KRFSLLMCGIWMVFLYMLPFVI---LKQEYHLVHSEITTCDDVVDACESPSSFRFY 277
Db 144 RIEHAKGVCIWFVILVFAQTLLPLINPMKOE-----AERITCMEYFNFEETKS--LPWI 196
Qy 278 FVSLAFFGFLPFVFIIFCYTTLTHK-----LKSODRIWLGYIKAVLLIVITIC 328
Db 197 LIGACFIGVVLPLIIILICYSQICKLFRPTAKQNPTEKSGVKKAKN---TVCLMNFNCMD 256
Qy 329 FAPTNIIIVH-----HANY-----YYHNTD-SLYFMYLIALCLGSLNSCLDPLFYFMS 377
Db 257 FTPYHVAIQHMVKRLRFSNFLECSQRHSFOISLHF-----TVCLMNFNCMDPIYFFAC 312
Qy 378 K 378
Db 313 K 313

RESULT 10

US-10-251-385-206
; Sequence 206, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496

; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 206
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-206

Query Match 20.1%; Score 428.5; DB 9; Length 361;
Best Local Similarity 31.1%; Pred. No. 1.7e-29;
Matches 96; Conservative 64; Mismatches 98; Indels 51; Gaps 10;
Qy 102 YLRSSLSTQVIPAIIYLLFVVGVPSNIVTLWKLRLTKSI-SLVIFHTNLAIADLLFCVT 160
Db 24 YAHSTARIWPLHYSLVFIIGLVGNLLALVIVQNRKKINSTLYSTNLVSDILFTTA 83
Qy 161 LPFKIAYHLNGNNWVGEVNCRIITVVFYGNMYCAIILTCMGINRYLATAHPTFYOKLP 220
Db 84 LPTRIAIYAMGFDWRIGDLCRITALVFYINTYAGVNFMTCLSIDRFAIVVHPLRYNKK 143
Qy 221 KRFSLLMCGIWMVFLYMLPFVI---LKQEYHLVHSEITTCDDVVDACESPSSFRFY 277
Db 144 RIEHAKGVCIWFVILVFAQTLLPLINPMKOE-----AERITCMEYFNFEETKS--LPWI 196
Qy 278 FVSLAFFGFLPFVFIIFCYTTLTHK-----LKSODRIWLGYIKAVLL 320
Db 197 LIGACFIGVVLPLIIILICYSQICKLFRPTAKQNPTEKSGVKKAKN---TVCLMNFNCMD 248
Qy 321 ILVIFTCFAPTNIIIVH-----HANY-----YYHNTD-SLYFMYLIALCLGSLNSCLD 369
Db 249 IIVVFLCTPTPYHVAIQHMVKRLRFSNFLECSQRHSFOISLHF-----TVCLMNFNCMD 304
Qy 370 PELYFVMSK 378
Db 305 PFYFFACK 313

RESULT 11

US-09-827-937A-17
; Sequence 17, Application US/09827937A
; Patent No. US20020052043A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human G-Protein Coupled Receptors
; FILE REFERENCE: 1488.1220003
; CURRENT APPLICATION NUMBER: US/09/827,937A
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 08/852,824
; PRIOR FILING DATE: 1997-05-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-827-937A-17

Query Match 20.0%; Score 426.5; DB 10; Length 348;
Best Local Similarity 31.6%; Pred. No. 2.5e-29;
Matches 95; Conservative 63; Mismatches 108; Indels 35; Gaps 9;
Qy 102 YLRSSLSTQVIPAIIYLLFVVGVPSNIVTLWKLRLTKSI-SLVIFHTNLAIADLLFCVT 160
Db 21 YAHSTARIWPLHYSLVFIIGLVGNLLALVIVQNRKKINSTLYSTNLVSDILFTTA 80
Qy 161 LPFKIAYHLNGNNWVGEVNCRIITVVFYGNMYCAIILTCMGINRYLATAHPTFYOKLP 220
Db 81 LPTRIAIYAMGFDWRIGDLCRITALVFYINTYAGVNFMTCLSIDRFAIVVHPLRYNKK 140
Qy 221 KRFSLLMCGIWMVFLYMLPFVI---LKQEYHLVHSEITTCDDVVDACESPSSFRFY 277

Db 141 RIEHAKGVCIFVWILVFAQTLLPLLNPMKQE-----AERITCMEYFNFEETKS--LPWI 193
Qy 278 FVSLARFGFLPFVILIFCVTTLLHK-----LKSRIWLGVIKAVLLILVIFTIC 328
Db 194 LLGACFIGYVLPFLIIKICISQICCKFRTRAKQNPLTKSGVKNKALNTIILIVVFLC 253
Qy 329 FAPNTIILVIH-----HANY-----YVHTND-SLYFMYLIALCIGLSNCLDPLFLVMS 377
Db 254 FTPVHVALIQHMWIKKLRFNSFLECSQRHSQISLHF-----TVCLMNFNCMDPFIYFAC 309
Qy 378 K 378
Db 310 K 310

RESULT 12
US-10-092-135-3
; Sequence 3, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
; TITLE OF INVENTION: HGPBMT27
; FILE REFERENCE: D0134.NP
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US/10/092,135
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 60/278,983
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 362
; TYPE: PRT
; ORGANISM: GALLUS GALLUS
US-10-092-135-3

Query Match 18.7%; Score 399; DB 9; Length 362;
Best Local Similarity 28.3%; Pred. No. 6.5e-27;
Matches 91; Conservative 63; Mismatches 122; Indels 46; Gaps 10;

Qy 73 GWTGATTITKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLW 132
Db 21 GWAAGNATTKC-----SLTKTGFQFYVLPVTVYILVFTIGFLGNSVAIW 63

Qy 133 KLSLRK---SISLVIFHTNLAIDLLFCVTLPPFKIAYHLNGNNWVGEVMCRTITTVFY 189
Db 64 MFVFMRPWSGISVYMF--NLALADFLYVLTLPALIFYFNKTDWIFGDMCKLQRFIFH 121

Qy 190 GNMVCAILILTCMGINRYLATAHPTVQKLPKRSFSLLMCGIWMVVMVFLYMLPFVILKQE 249
Db 122 VNLGYSILFCTCISVHRVYTGVVHPLKSLGRLKKNNAVYVSSLVWALVAVIAP-ILFYSG 180

Qy 250 YHLVHSEITTCDDVDACESPSSFRFYFVSIAFFGLIPFVILFVVIIFCYTTLIHLKSKD- 308
Db 181 TGVRRKNTICYDTT-ADEYLRSY-FVYSMCTTVFMFCIPFVILGCGYGLIVKALYKDL 238

Qy 309 -----RIWLGVIKAVLLILVIFTICFAPNTIILVIH-HANYYYHN-----TDSLYF 353
Db 239 DNSPLRRKSIYL-----VVIIVTFVAVSYLPFHVMKTLNRLARLDFQTPQMCAFNKDYA 293

Qy 354 MYLIALCLGSLNSCLDPLFLYFV 375
Db 294 TYQVTRGLASLNSCVDPIYFL 315

RESULT 13
US-09-779-679-28
; Sequence 28, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J

; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1e1 Proteins and Nucleic Acids Encoding the Sa
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779,679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180,929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-779-679-28

Query Match 18.7%; Score 399; DB 9; Length 362;
Best Local Similarity 28.3%; Pred. No. 6.5e-27;
Matches 91; Conservative 63; Mismatches 122; Indels 46; Gaps 10;

Qy 73 GWTGATTITKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLW 132
Db 21 GWAAGNATTKC-----SLTKTGFQFYVLPVTVYILVFTIGFLGNSVAIW 63

Qy 133 KLSLRK---SISLVIFHTNLAIDLLFCVTLPPFKIAYHLNGNNWVGEVMCRTITTVFY 189
Db 64 MFVFMRPWSGISVYMF--NLALADFLYVLTLPALIFYFNKTDWIFGDMCKLQRFIFH 121

Qy 190 GNMVCAILILTCMGINRYLATAHPTVQKLPKRSFSLLMCGIWMVVMVFLYMLPFVILKQE 249
Db 122 VNLGYSILFCTCISVHRVYTGVVHPLKSLGRLKKNNAVYVSSLVWALVAVIAP-ILFYSG 180

Qy 250 YHLVHSEITTCDDVDACESPSSFRFYFVSIAFFGLIPFVILFVVIIFCYTTLIHLKSKD- 308
Db 181 TGVRRKNTICYDTT-ADEYLRSY-FVYSMCTTVFMFCIPFVILGCGYGLIVKALYKDL 238

Qy 309 -----RIWLGVIKAVLLILVIFTICFAPNTIILVIH-HANYYYHN-----TDSLYF 353
Db 239 DNSPLRRKSIYL-----VVIIVTFVAVSYLPFHVMKTLNRLARLDFQTPQMCAFNKDYA 293

Qy 354 MYLIALCLGSLNSCLDPLFLYFV 375
Db 294 TYQVTRGLASLNSCVDPIYFL 315

RESULT 14

```
US-09-828-478-6
; Sequence 6, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match      18.7%; Score 398.5; DB 9; Length 367;
Best Local Similarity 28.3%; Pred. No. 7.3e-27;
Matches 98; Conservative 60; Mismatches 136; Indels 55; Gaps 9;

QY 42 SDNSAKPTLTIKSPNG---GPQNTFEFPPLSDIEGWTGATTTIKAECPDSISTLHVNNA 98
Db 23 SDSS-----QSMNGLEVAPPGLITNFSLATAE-----QCQE-----54
QY 99 TIGVLRSSLTQVIPAIYILFVVGVPSNIVTLWKLRLTKS-ISLVIFHTNLAIAIDLFL 157
Db 55 -----TPLENMLFASPYLLDFILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSC 108
QY 158 CVTLFPFKIAVHLNGNNWVFGVMCRITTVVFGYNNMYCAIILITCMGINRYLATAPHTYQ 217
Db 109 VLVLPTRLVVHFSGNHWPFGIEACRLTGFLFYLNMYASIFLTICISADRFIAIVHPVKSL 168
QY 218 KLPRKSFSLMCGIWMVVMFLYMLPFPVILKQYHLVHSEITTCCHDVVDACESPSPFRFY 277
Db 169 KLRPLVAHLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQYREKASHHA 221
QY 278 FVSLAPFGFLIPFVVIIFCYTTLIHKLK-----KDRWLGVKAVLILVIFTCFAPT 333
Db 222 LVSLA-VAFTFPFITVTCYLLIIRSLRQGLRVEKRLKTRAVRMIAIVLAIFLVCFVPYH 280
QY 334 IILVIHANYHTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 281 VNRSVVLHYRSHGASCATORILALANRITSLTSLNGALDPIYFFVAE 330

RESULT 15
US-09-828-478-4
; Sequence 4, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-4

Query Match      18.7%; Score 398.5; DB 9; Length 339;
Best Local Similarity 31.2%; Pred. No. 1.3e-26;
Matches 88; Conservative 53; Mismatches 123; Indels 18; Gaps 5;

QY 107 LSTQVIPAIYILFVVGVPSNIVTLWKLRLTKS-ISLVIFHTNLAIAIDLFLFCVTLPFKI 165
Db 29 LENMLFASPYLLDFILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSCVLLVPLTRL 88
QY 166 AYHLNGNNWVFGVMCRITTVVFGYNNMYCAIILITCMGINRYLATAPHTYQKLPKRSFS 225
Db 89 VVHFSGNHWPFGIEACRLTGFLFYLNMYASIFLTICISADRFIAIVHPVKSLKLRPLYA 148
QY 226 LLMCGIWMVVMFLYMLPFPVILKQYHLVHSEITTCCHDVVDACESPSPFRFYFVSLA 285
Db 149 HLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQYREKASHHALVSLA-VA 200
QY 286 FLIPFVVIIFCYTTLIHKLK-----KDRWLGVKAVLILVIFTCFAPTNIILVIHHA 341
Db 201 FTFFPFITVTCYLLIIRSLRQGLRVEKRLKTRAVRMIAIVLAIFLVCFVPYHVNRSVYL 260
QY 342 NYYYHNTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 261 HYRSHGASCATORILALANRITSLTSLNGALDPIYFFVAE 302

RESULT 16
US-10-251-385-32
; Sequence 32, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human
; TITLE OF INVENTION: Protein-Coupled
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 32
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-32

Query Match      18.5%; Score 395; DB 9; Length 339;
Best Local Similarity 31.2%; Pred. No. 1.3e-26;
Matches 88; Conservative 53; Mismatches 123; Indels 18; Gaps 5;

QY 107 LSTQVIPAIYILFVVGVPSNIVTLWKLRLTKS-ISLVIFHTNLAIAIDLFLFCVTLPFKI 165
Db 29 LENMLFASPYLLDFILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSCVLLVPLTRL 88
QY 166 AYHLNGNNWVFGVMCRITTVVFGYNNMYCAIILITCMGINRYLATAPHTYQKLPKRSFS 225
Db 89 VVHFSGNHWPFGIEACRLTGFLFYLNMYASIFLTICISADRFIAIVHPVKSLKLRPLYA 148
QY 226 LLMCGIWMVVMFLYMLPFPVILKQYHLVHSEITTCCHDVVDACESPSPFRFYFVSLA 285
Db 149 HLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQYREKASHHALVSLA-VA 200
QY 286 FLIPFVVIIFCYTTLIHKLK-----KDRWLGVKAVLILVIFTCFAPTNIILVIHHA 341
Db 201 FTFFPFITVTCYLLIIRSLRQGLRVEKRLKTRAVRMIAIVLAIFLVCFVPYHVNRSVYL 260
QY 342 NYYYHNTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 261 HYRSHGASCATORILALANRITSLTSLNGALDPIYFFVAE 302
```

```
US-09-828-478-6
; Sequence 6, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match      18.7%; Score 398.5; DB 9; Length 367;
Best Local Similarity 28.3%; Pred. No. 7.3e-27;
Matches 98; Conservative 60; Mismatches 136; Indels 55; Gaps 9;

QY 42 SDNSAKPTLTIKSPNG---GPQNTFEFPPLSDIEGWTGATTTIKAECPDSISTLHVNNA 98
Db 23 SDSS-----QSMNGLEVAPPGLITNFSLATAE-----QCQE-----54
QY 99 TIGVLRSSLTQVIPAIYILFVVGVPSNIVTLWKLRLTKS-ISLVIFHTNLAIAIDLFL 157
Db 55 -----TPLENMLFASPYLLDFILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSC 108
QY 158 CVTLFPFKIAVHLNGNNWVFGVMCRITTVVFGYNNMYCAIILITCMGINRYLATAPHTYQ 217
Db 109 VLVLPTRLVVHFSGNHWPFGIEACRLTGFLFYLNMYASIFLTICISADRFIAIVHPVKSL 168
QY 218 KLPRKSFSLMCGIWMVVMFLYMLPFPVILKQYHLVHSEITTCCHDVVDACESPSPFRFY 277
Db 169 KLRPLVAHLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQYREKASHHA 221
QY 278 FVSLAPFGFLIPFVVIIFCYTTLIHKLK-----KDRWLGVKAVLILVIFTCFAPT 333
Db 222 LVSLA-VAFTFPFITVTCYLLIIRSLRQGLRVEKRLKTRAVRMIAIVLAIFLVCFVPYH 280
QY 334 IILVIHANYHTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 281 VNRSVVLHYRSHGASCATORILALANRITSLTSLNGALDPIYFFVAE 330

RESULT 15
US-09-828-478-4
; Sequence 4, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-4

Query Match      18.5%; Score 395; DB 9; Length 339;
Best Local Similarity 31.2%; Pred. No. 1.3e-26;
Matches 88; Conservative 53; Mismatches 123; Indels 18; Gaps 5;

QY 107 LSTQVIPAIYILFVVGVPSNIVTLWKLRLTKS-ISLVIFHTNLAIAIDLFLFCVTLPFKI 165
Db 29 LENMLFASPYLLDFILALVGNLTALWLFIRDHKSGTTPANVFLMHLAVADLSCVLLVPLTRL 88
QY 166 AYHLNGNNWVFGVMCRITTVVFGYNNMYCAIILITCMGINRYLATAPHTYQKLPKRSFS 225
Db 89 VVHFSGNHWPFGIEACRLTGFLFYLNMYASIFLTICISADRFIAIVHPVKSLKLRPLYA 148
QY 226 LLMCGIWMVVMFLYMLPFPVILKQYHLVHSEITTCCHDVVDACESPSPFRFYFVSLA 285
Db 149 HLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQYREKASHHALVSLA-VA 200
QY 286 FLIPFVVIIFCYTTLIHKLK-----KDRWLGVKAVLILVIFTCFAPTNIILVIHHA 341
Db 201 FTFFPFITVTCYLLIIRSLRQGLRVEKRLKTRAVRMIAIVLAIFLVCFVPYHVNRSVYL 260
QY 342 NYYYHNTDS-----LYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 261 HYRSHGASCATORILALANRITSLTSLNGALDPIYFFVAE 302
```

RESULT 17

US-09-848-889-12

; Sequence 12, Application US/09848889

; Patent No. US2002002555A1

; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice

; APPLICANT: Guegler, Karl J.

; APPLICANT: Cheng, Muzong

; TITLE OF INVENTION: GPCR DIAGNOSTIC FOR BRAIN CANCER

; FILE REFERENCE: PC-0042 CIP

; CURRENT APPLICATION NUMBER: US/09/848.889

; CURRENT FILING DATE: 2001-05-03

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PERL Program

; SEQ ID NO 12

; LENGTH: 339

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US2002002555A1 9992700

US-09-848-889-12

Query Match 18.5%; Score 395; DB 10; Length 339;

Best Local Similarity 31.2%; Pred. No. 1.3e-26;

Matches 88; Conservative 53; Mismatches 123; Indels 18; Gaps 5;

Qy 107 LSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRTKS-ISLVIFHTNLAIADLLFCVTLPLFKI 165

Db 29 LENMLFASFYLLDFILALVGNLTALWLFIRDHKSQTPANVFLMHLAVADLSCVLPTRL 88

Qy 166 AYHLGNWVFGVEMCRITTVVYGNMYCAIILITCMGINRYLATAHPFTYOKLPKRSFS 225

Db 89 VYHFSGNHWPGEIACRLTGLFYLNNMYASIFLTICISADRFALVHPVKSLLRPLA 148

Qy 226 LLMCGIYVMVFLYMLPFVILKQYHLVHSEITTCCHDVVDACESPSSFRFYFVSLAFFG 285

Db 149 HLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQLYREKASHHALVSLA-VA 200

Qy 286 FLIPFVIIICYTTLIHLKLS-----KDRIMLGYIKAVLLILVIFTICFAPTNIILVIHHA 341

Db 201 FTFFPITVTCLLIIRSLRQGLRVEKRLTKAVRMIAIIVLFLVCFVPHVNRSVYVL 260

Qy 342 NYIYHNTDS-----LYFMYLIALCLGSLNSCLDPLFYFVMSK 378

Db 261 HYRSHGASCATORILALANRITSLTSLNGALDPIIMYFFVAE 302

RESULT 18

US-09-788-133-2

; Sequence 2, Application US/09788133

; Patent No. US20020052001A1

; GENERAL INFORMATION:

; APPLICANT: GLAXO GROUP LTD

; APPLICANT: ASSAY

; TITLE OF INVENTION: ASSAY

; FILE REFERENCE: P79011

; CURRENT APPLICATION NUMBER: US/09/788.133

; CURRENT FILING DATE: 2001-02-20

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 2

; LENGTH: 339

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-788-133-2

Query Match 18.5%; Score 395; DB 10; Length 339;

Best Local Similarity 31.2%; Pred. No. 1.3e-26;

Matches 88; Conservative 53; Mismatches 123; Indels 18; Gaps 5;

Qy 107 LSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRTKS-ISLVIFHTNLAIADLLFCVTLPLFKI 165

Db 29 LENMLFASFYLLDFILALVGNLTALWLFIRDHKSQTPANVFLMHLAVADLSCVLPTRL 88

Qy 166 AYHLGNWVFGVEMCRITTVVYGNMYCAIILITCMGINRYLATAHPFTYOKLPKRSFS 225

Db 89 VYHFSGNHWPGEIACRLTGLFYLNNMYASIFLTICISADRFALVHPVKSLLRPLA 148

Qy 226 LLMCGIYVMVFLYMLPFVILKQYHLVHSEITTCCHDVVDACESPSSFRFYFVSLAFFG 285

Db 149 HLACAFLLWVAVAMAPLLVSPQ-----TVQTNHTVCLQLYREKASHHALVSLA-VA 200

Qy 286 FLIPFVIIICYTTLIHLKLS-----KDRIMLGYIKAVLLILVIFTICFAPTNIILVIHHA 341

Db 201 FTFFPITVTCLLIIRSLRQGLRVEKRLTKAVRMIAIIVLFLVCFVPHVNRSVYVL 260

Qy 342 NYIYHNTDS-----LYFMYLIALCLGSLNSCLDPLFYFVMSK 378

Db 261 HYRSHGASCATORILALANRITSLTSLNGALDPIIMYFFVAE 302

RESULT 19

US-10-092-135-4

; Sequence 4, Application US/10092135

; Publication No. US20030054374A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR

; TITLE OF INVENTION: HGPBMY27

; FILE REFERENCE: D0134.NP

; CURRENT APPLICATION NUMBER: US/10/092.135

; CURRENT FILING DATE: 2002-03-06

; PRIOR APPLICATION NUMBER: US 60/273,808

; PRIOR FILING DATE: 2001-03-07

; PRIOR APPLICATION NUMBER: US 60/278,983

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 75

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 4

; LENGTH: 362

; TYPE: PRT

; ORGANISM: MELEAGRIS GALLOPAVO

US-10-092-135-4

Query Match 18.5%; Score 395; DB 9; Length 362;

Best Local Similarity 28.0%; Pred. No. 1.5e-26;

Matches 90; Conservative 64; Mismatches 122; Indels 46; Gaps 10;

Qy 73 GMTGATTTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLW 132

Db 21 GWAAGNASTKC-----SLTKTGQFYLLPTVYILVFTITGLGNSVAIW 63

Qy 133 KLSLRTK---SISLVIFHTNLAIADLLFCVTLPLFKIAYHLGNWVFGVEMCRITTVFY 189

Db 64 MFVFMHMPWSGISVYMF--NLALADFLYVLTLPALIFYFNFNTDITFGDMCKLQRFIFH 121

Qy 190 GNMCAIILITCMGINRYLATAHPFTYOKLPKRSFSLMCGIYVMVFLYMLPFVILKQE 249

Db 122 VNLGSIILFTLCISVHRYTGCVVHPLKSLGRLKKNAVYSSLVWALVAVIAP-ILFYSG 180

Qy 250 YHLVHSEITTCCHDVVDACESPSSFRFYFVSLAFFGFLIPFVIIICYTTLIHLKSKD- 308

Db 181 TGVRRNKITTCYDIT-ADVLRSY-FVYSMCTTVFMFCIPFVILGCGYGLIVKALIVKDL 238

Qy 309 -----RIWLGYIKAVLLILVIFTICFAPTNIILVIH-HANYYYHN-----TSLYF 353

Db 239 DNSPLRRKSIYL-----VIIIVTFVAVSYLPPHVMKTLNLRARLDFQTPMCAMFNKDYA 293

Qy 354 MYLIALCLGSLNSCLDPLFYFV 375

Db 294 TYQVTRGLASLNSCVDPILYFL 315

RESULT 20


```
US-09-826-791-6
; Sequence 6, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1el Polypeptide
; FILE REFERENCE: PC10914ADAM
; CURRENT APPLICATION NUMBER: US/09/826,791
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 0008504.3
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/198,367
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-791-6

Query Match 17.7%; Score 378.5; DB 10; Length 346;
Best Local Similarity 28.9%; Pred. No. 3.8e-25;
Matches 88; Conservative 63; Mismatches 121; Indels 33; Gaps 8;

QY 86 PEDSISTLHVNNTATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLWK-LSLRTKTSISLV 144
Db 19 PNGTFSNNNSRNCIT-----ENFKREFFPIVLIIFFWGLNGLSIYVFLQPYKKSTSVN 74
QY 145 IFHTNLAIADLLFCVTLFPFKIAYHLNGNNWVGEVWMCRTITTVFYGNMYCAIILTCMGI 204
Db 75 VFMLNLAIADLLFISTLPFRADYILRGSNWIFGDLACRIMSYSLYVNMYSIYFLTVLSV 134
QY 205 NRYLATAPHTYQKLPKRSFSLMCGIWMVVMFLYMLPFVILKQEHVHLVHSEITTCDDVV 264
Db 135 VRFLAMVHPFRLHLVTSIRSAWILCGIWIILI---MASSIMLLDSDGSEQNGSVTSCLEL- 190
QY 265 DACESPSSFRFYFVSL-----AFFGFLIPFVIIIFCYVTTLIHKLKLS---KDRILWL 313
Db 191 -----NLYKIAKLOTNMYIALVVGCLLPFFTLISCYLLIIRVLLKVEVPESGLRVS 241
QY 314 YIKA---VLLILVIFTICFAPTNIIIVHANYHH-NTDSLYFMYLIAICLGLSLSCLD 369
Db 242 HRKALTTIITLIIFLCFLPYHTLRTVHLTWKVGCKDRHLKALVITLALAAANACFN 301
QY 370 PFLYF 374
Db 302 PLYY 306

RESULT 28
US-09-866-230-7
; Sequence 7, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Ma
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 7
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230-7

Query Match 17.7%; Score 378.5; DB 10; Length 346;
Best Local Similarity 28.9%; Pred. No. 3.8e-25;
Matches 88; Conservative 63; Mismatches 121; Indels 33; Gaps 8;

QY 86 PEDSISTLHVNNTATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLWK-LSLRTKTSISLV 144
Db 19 PNGTFSNNNSRNCIT-----ENFKREFFPIVLIIFFWGLNGLSIYVFLQPYKKSTSVN 74
QY 145 IFHTNLAIADLLFCVTLFPFKIAYHLNGNNWVGEVWMCRTITTVFYGNMYCAIILTCMGI 204
Db 75 VFMLNLAIADLLFISTLPFRADYILRGSNWIFGDLACRIMSYSLYVNMYSIYFLTVLSV 134
QY 205 NRYLATAPHTYQKLPKRSFSLMCGIWMVVMFLYMLPFVILKQEHVHLVHSEITTCDDVV 264
Db 135 VRFLAMVHPFRLHLVTSIRSAWILCGIWIILI---MASSIMLLDSDGSEQNGSVTSCLEL- 190
QY 265 DACESPSSFRFYFVSL-----AFFGFLIPFVIIIFCYVTTLIHKLKLS---KDRILWL 313
Db 191 -----NLYKIAKLOTNMYIALVVGCLLPFFTLISCYLLIIRVLLKVEVPESGLRVS 241
QY 314 YIKA---VLLILVIFTICFAPTNIIIVHANYHH-NTDSLYFMYLIAICLGLSLSCLD 369
Db 242 HRKALTTIITLIIFLCFLPYHTLRTVHLTWKVGCKDRHLKALVITLALAAANACFN 301
QY 370 PFLYF 374
Db 302 PLYY 306

RESULT 28
US-09-866-230-7
; Sequence 7, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 7
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230-7

Query Match 17.7%; Score 378.5; DB 10; Length 346;
Best Local Similarity 28.9%; Pred. No. 3.8e-25;
Matches 88; Conservative 63; Mismatches 121; Indels 33; Gaps 8;
```

```
RESULT 30
US-09-828-478-5
; Sequence 5, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-5

Query Match      17.7%; Score 378; DB 9; Length 337;
Best Local Similarity 32.6%; Pred. No. 4e-25;
Matches 98; Conservative 55; Mismatches 128; Indels 20; Gaps 9;

QY 88 DSISTLHVNNTATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT--KSIISLVI 145
Db 2 DETGNLTSSATCHDTIDDFRNQVYSTLYSMISVGVFGFGFVLYVL-IKTYHKKSAPQV 60
QY 146 FHTNLAIADLLFCVTLPEFKIAYHLNGNNWVFGVMCRITTVFYGNMYCAILITLTCMGIN 205
Db 61 YMINLAVADLLCVCTPLRVRVYVYHKGIMLFGDFCLRLSTYALYVNLVYCSIFFTWMSFF 120
QY 206 RYLATAHPFTYQKLPRKSFSLMCGIWMVVMFLYMLPFVILKOEYHLVHSEIITTCHEVD 265
Db 121 RCIAIVFPQVNLVTKARFVCGVIGWIFVILTSSPFLMAKPKQD--EKNTKCFEPPQ 178
QY 266 ACESPSFRFYFVSLAFGLIPFVILFCYTTLIHLK--KSKDRWLWGIYKAVLLILV 323
Db 179 DNQTKNHLVHLHYVSL-FVGFIIPFVILVVCYTMIIITLLKSKMKNLSSHKAIGMIMV 237
QY 324 I---FTICFAPTNIILVIHANYVHN---TDSLYFM---YLIALCLGSLNSCLDPFLY 373
Db 238 VTAFLVSPFPHIQRTIH--LHFLHNETKPCDSVLRMQKSVVITLSLAASNCDFDPLLY 295
QY 374 F 374
Db 296 F 296

Query Match      17.7%; Score 378; DB 9; Length 337;
Best Local Similarity 32.6%; Pred. No. 4e-25;
Matches 98; Conservative 55; Mismatches 128; Indels 20; Gaps 9;

QY 88 DSISTLHVNNTATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT--KSIISLVI 145
Db 2 DETGNLTSSATCHDTIDDFRNQVYSTLYSMISVGVFGFGFVLYVL-IKTYHKKSAPQV 60
QY 146 FHTNLAIADLLFCVTLPEFKIAYHLNGNNWVFGVMCRITTVFYGNMYCAILITLTCMGIN 205
Db 61 YMINLAVADLLCVCTPLRVRVYVYHKGIMLFGDFCLRLSTYALYVNLVYCSIFFTWMSFF 120
QY 206 RYLATAHPFTYQKLPRKSFSLMCGIWMVVMFLYMLPFVILKOEYHLVHSEIITTCHEVD 265
Db 121 RCIAIVFPQVNLVTKARFVCGVIGWIFVILTSSPFLMAKPKQD--EKNTKCFEPPQ 178
QY 266 ACESPSFRFYFVSLAFGLIPFVILFCYTTLIHLK--KSKDRWLWGIYKAVLLILV 323
Db 179 DNQTKNHLVHLHYVSL-FVGFIIPFVILVVCYTMIIITLLKSKMKNLSSHKAIGMIMV 237
QY 324 I---FTICFAPTNIILVIHANYVHN---TDSLYFM---YLIALCLGSLNSCLDPFLY 373
Db 238 VTAFLVSPFPHIQRTIH--LHFLHNETKPCDSVLRMQKSVVITLSLAASNCDFDPLLY 295
QY 374 F 374
Db 296 F 296

RESULT 31
US-09-866-230-8
; Sequence 8, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 8
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-230-8
```

```
Query Match      17.7%; Score 378; DB 10; Length 337;
Best Local Similarity 32.6%; Pred. No. 4e-25;
Matches 98; Conservative 55; Mismatches 128; Indels 20; Gaps 9;

QY 88 DSISTLHVNNTATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT--KSIISLVI 145
Db 2 DETGNLTSSATCHDTIDDFRNQVYSTLYSMISVGVFGFGFVLYVL-IKTYHKKSAPQV 60
QY 146 FHTNLAIADLLFCVTLPEFKIAYHLNGNNWVFGVMCRITTVFYGNMYCAILITLTCMGIN 205
Db 61 YMINLAVADLLCVCTPLRVRVYVYHKGIMLFGDFCLRLSTYALYVNLVYCSIFFTWMSFF 120
QY 206 RYLATAHPFTYQKLPRKSFSLMCGIWMVVMFLYMLPFVILKOEYHLVHSEIITTCHEVD 265
Db 121 RCIAIVFPQVNLVTKARFVCGVIGWIFVILTSSPFLMAKPKQD--EKNTKCFEPPQ 178
QY 266 ACESPSFRFYFVSLAFGLIPFVILFCYTTLIHLK--KSKDRWLWGIYKAVLLILV 323
Db 179 DNQTKNHLVHLHYVSL-FVGFIIPFVILVVCYTMIIITLLKSKMKNLSSHKAIGMIMV 237
QY 324 I---FTICFAPTNIILVIHANYVHN---TDSLYFM---YLIALCLGSLNSCLDPFLY 373
Db 238 VTAFLVSPFPHIQRTIH--LHFLHNETKPCDSVLRMQKSVVITLSLAASNCDFDPLLY 295
QY 374 F 374
Db 296 F 296

RESULT 32
US-10-092-135-7
; Sequence 7, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: D0134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 7
; LENGTH: 373
; TYPE: PRT
; ORGANISM: RATTUS NORVEGICUS
US-10-092-135-7

Query Match      17.7%; Score 377.5; DB 9; Length 373;
Best Local Similarity 29.3%; Pred. No. 5e-25;
Matches 86; Conservative 65; Mismatches 110; Indels 33; Gaps 9;

QY 103 LRSSLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT--SISLVIFHTNLAIADLLFCV 159
Db 45 IKTGFGFYLPVAVIILVFIIFGLGNSVAIWMFVHMKPWSGISVYWF--NLADLFLYL 102
QY 160 TLPFKIAYHLNGNNWVFGVMCRITTVFYGNMYCAILITLTCMGINRYLATAHPFTYQKL 219
Db 103 TLPALIFYFNKTDWIFGDVCMCKLQRFIFHVNLYGSILFLTCSIAHRYSGVYVPLKSLGR 162
QY 220 PKRSFSLMCGIWMVVMFLYMLPFVILKOEYHLVHSEIITTCHEVDACSPSFRFYFV 279
Db 163 LKKNAIYVSVLWLIWVAISP-ILFYSGTGIRKNTVTCYD---STSDYLRSYFIY 217
QY 280 SL--AFEGFLIPFVILFCYTTLIHLKSKD-----RIWGIYKAVLLILVIFTC 328
Db 218 SMCTTVAMFCIPLVLILGCVGLIVRALIYKDLNDSPLRRKSIYL-----VIIIVTFAVS 272
```



```
Db 179 DNQTKHVLVHLVSL-FVGGIIPFVIIIVCYTWIILTLKSKMKNLSHKKAIGMIV 237
Qy 324 I---FTICFAPTNIILVIHANYHHN---TDSLIFM---YLIALCLGSLNSCLDPLFY 373
Db 238 VTAFLVSPFYHQIHTI--LHFLHNETKPCDVLRMQKSVVITLSLAASNCDFDPLLY 295
Qy 374 F 374
Db 296 F 296

RESULT 35
US-10-092-135-6
; Sequence 6, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
; FILE REFERENCE: D0134.NP
; CURRENT APPLICATION NUMBER: US/10/092.135
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 373
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-092-135-6
Query Match 17.6%; Score 376.5; DB 9; Length 373;
Best Local Similarity 27.0%; Pred. No. 6.2e-25;
Matches 95; Conservative 72; Mismatches 130; Indels 55; Gaps 12;

Qy 46 AKPTLTITKSFNGGQNTFEFFPLSDIEGTGATTIKAECPEDSISTLHVNNATIGYLRL 104
Db 8 AVPNGTDAFLAGPSS-----WGNSTVA-----STAAVSSSFKCALTK 46

Qy 105 SSLSQTVPIPAIYILLFVVGVPNSIVTLWLSLRTK---SISLVIFHTNLAIADLLFCVTL 161
Db 47 TGFQFYVLPVAVIYLVFIIGLNSVAIWMFVFMKPSGIVYMF--NLALADFLYVLT 104

Qy 162 PFKIAYHLNGNNWVGEVCMCRITTVVFGNMYCAIILITCMGINRYLATAHPTTYQKLPK 221
Db 105 PALIFYFYNKTDWIFGDAMCKLQRFIFHVNLYGSILFTICSAHRYSGVVYVPLKSLGRK 164

Qy 222 RSFLLMCGIWMVFLYMLPFVILKQYHLVHSEITTDHVDVADCESSPSFRFYFVSL 281
Db 165 KNAICISVLWLIIVVAISP-ILFYSGTGVRKNKTICTYDTT-----SDEYLSRYFIYSM 219

Qy 282 --AFFGFLIPFVIIFCYTTLIHKLKSKD-----RIWLGVIKAVLLILVIFTCIFA 330
Db 220 CTTVAMFCVPLVLIGCYLVIRALYKLDNSPLRKSIVL-----VIIVLTVFAVSYI 274

Qy 331 PTNIILVIH-HANYHHN-----TDSLIFMYLIALCLGSLNSCLDPLFYFV 375
Db 275 PFVWMTNLRARLDFTPAWCAFNDRVYATYQVTRGLASLNSCVDPLVFL 326

RESULT 36
US-10-092-135-5
; Sequence 5, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
; FILE REFERENCE: D0134.NP
; CURRENT APPLICATION NUMBER: US/10/092.135
```

```
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 373
; TYPE: PRT
; ORGANISM: BOS TAURUS
US-10-092-135-5
Query Match 17.4%; Score 372; DB 9; Length 373;
Best Local Similarity 27.8%; Pred. No. 1.5e-24;
Matches 94; Conservative 63; Mismatches 139; Indels 42; Gaps 10;

Qy 59 PQNTFEFPLSDIEGTGATTIKAECPEDSISTLHVNNATIGYLRSSLSQTQVIPAIIYL 118
Db 10 PNGTDTAFLADPGSPMGNSTVTSTAASVAFKCALTKTGFQFYFL-----PAVYIL 60

Qy 119 LFWVGVPNSIVTLWLSLRTK---SISLVIFHTNLAIADLLFCVTLFPFKIAYHLNGNNV 175
Db 61 VFIIGLNSVAIWMFVFMKPSGIVYMF--NLALADFLYVLTLPALIFYFYNKTDWI 118

Qy 176 FGEVCMCRITTVVFGNMYCAIILITCMGINRYLATAHPTTYQKLPKRSFSLMCGIWMV 235
Db 119 FGDAMCKLQRFIFHVNLYGSILFTICSAHRYSGVVYVPLKSLGRKKNAVYISVLWLI 178

Qy 236 VFYMLPFVILKQYHLVHSEITTDHVDVADCESSPSFRFYFVSL--AFFGFLIPFVII 293
Db 179 VVVGISP-ILFYSGTGIRKNKTICTYDTT-----SDEYLSRYFIYSCTTVAMFCVPLVLI 233

Qy 294 IFCYTTLIHKLKSKD-----RIWLGVIKAVLLILVIFTCIFAPTNIILVIH-HANY 343
Db 234 LGCYGLIVRALYKLDNSPLRKSIVL-----VIIVLTVFAVSYIPFVWMTNLRARL 288

Qy 344 YVHN-----TDSLIFMYLIALCLGSLNSCLDPLFYFV 375
Db 289 DFQTPMCAFNDRVYATYQVTRGLASLNSCVDPLVFL 326

RESULT 37
US-09-943-718-4
; Sequence 4, Application US/09943718
; Patent No. US20020103361A1
; GENERAL INFORMATION:
; APPLICANT: Hufine, Constance F.
; Rossi, Devora L.
; Capone, Myriam
; Hedrick, Joseph A.
; Vicari, Alain
; Gorman, Daniel M.
; Zlotnik, Albert
; TITLE OF INVENTION: Mammalian Chemokines; Receptors;
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/943,718
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/009,817
; FILING DATE: 20-APR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX058K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-943-718-4
Query Match 17.4%; Score 371.5; DB 10; Length 360;
Best Local Similarity 32.8%; Pred. No. 1.6e-24;
Matches 78; Conservative 48; Mismatches 93; Indels 19; Gaps 4;
QY 149 NLAIADLLFCVTELPFKIAYHLNGNNWVFGVMCRITTWVFGYNNMYCAILITCMGINRYL 208
DB 93 NLAVADLLALVLPRLAYHLRGQRPFGAACHRVATAALYGHMYGVSLLAAVSLDRYL 152
QY 209 ATAHPTFYQKLPKRSFSLMCGIVVMVFLPVLKQBYHLVHSEITTHCHVDVAC- 267
DB 153 ALVHPLRARALRGQRLTGLCLVAWLSAATLALPLTLHRQNFRL-APAIACC--VMWRCP 209
QY 268 -----ESPSSFRFYFVSLAFFGLIPFVILFCVTTLIHKLSKDRILGVIKAVL 319
DB 210 WLSTPTGERPSSAWLSWAASL-----PULAMGLCYGTTLRALAANGQRYSHALRLTA 262
QY 320 LILVIFTICFAPTNILVIHANYHYNTDSLYFMYLIALCLGSLNSCLDPFLFYVMS 377
DB 263 LVLFSAVASFTPSNVLVLVHVSNPSPANGNLYGAYVPSLALSTLNSCVDPFYIYVYS 320
RESULT 38
US-10-167-192-3
; Sequence 3, Application US/10167192
; Publication No. US20030040052A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; FILE OF INVENTION: No. US20030040052A1el G-Protein Coupled Receptors
; CURRENT APPLICATION NUMBER: US/10/167,192
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: US/09/420,187
; PRIOR FILING DATE: 1999-10-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-167-192-3
Query Match 17.3%; Score 370; DB 9; Length 337;
Best Local Similarity 32.2%; Pred. No. 2e-24;
Matches 97; Conservative 55; Mismatches 129; Indels 20; Gaps 9;
QY 88 DSISTLHVNNTATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT--KSLSLVI 145
DB 2 DETGNLTVSSATCHDTIDDFRNQVSTLYSMISVVGFNGFVLYVL-IKTYHKKSAPQV 60
QY 146 FHTNLAIADLLFCVTELPFKIAYHLNGNNWVFGVMCRITTWVFGYNNMYCAILITCMGIN 205
DB 61 YMINLAVADLLCVCTPLPVRVYVYVHKGIWLFGLCRSLTVALYNLYCSIFFMTAWSFF 120
QY 206 RYLATAHPFTYQKLPKRSFSLMCGIVVMVFLPVLKQBYHLVHSEITTHCHDVVD 265
```

```
DB 121 RCIAIVFPYQINLVTKKARFVCGIWMIFVILTSSPFLMAKPKD--EKNNTKCFEPQ 178
QY 266 ACESPSFRFYFVSVSLAFPGFLIPFVIIIFCVTTLIHKL--KSKDRILGVIKAVLLILV 323
DB 179 DNQTKNHLVLVHVSIVL-VGFIIPFVIIIVCYTMIITLLTKSKMKNLSHSHKKAIGMIV 237
QY 324 I---FTICFAPTNILVIHANYHYHN---TDSLYFM---YLIALCLGSLNSCLDPFLY 373
DB 238 VTAAFLVSFMPHIQRTIH--LHFLHNETKPCOSVLRMQKSVVITLSLAASNCDFPLLY 295
QY 374 F 374
DB 296 F 296
RESULT 39
US-10-023-775B-2
; Sequence 2, Application US/10023775B
; Publication No. US20030022282A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (EP,GB) only
; APPLICANT: Pfizer Inc. (US, JP, EB except GB)
; APPLICANT: Fidoock, Mark David
; TITLE OF INVENTION: No. US20030022282A1el Polypeptide
; FILE REFERENCE: PC10959AGPR
; CURRENT APPLICATION NUMBER: US/10/023,775B
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: GB 0030854.4
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,590
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/296,660
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0111031.1
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-775B-2
Query Match 17.3%; Score 369.5; DB 9; Length 337;
Best Local Similarity 29.8%; Pred. No. 2.2e-24;
Matches 87; Conservative 58; Mismatches 108; Indels 41; Gaps 8;
QY 107 LSTQVIPAIYILLFVVGVPNSIVTL-----WKLRLTKSISLVIFHTNLAIADLLF 157
DB 31 LKMHYLPVIYGIIFLVGFPFNNAVISTYIFKMRPWK-----SSTIIMNLACTDLY 82
QY 158 CVTLPPKIAHYHLNGNNWVFGVMCRITTWVFGYNNMYCAILITCMGINRYLATAHPFTYQ 217
DB 83 LTSPLPELIHYASGENWIFGDFMCKFRFSEHLYSSILFLFCFIFRYCVIHPMSCF 142
QY 218 KLPKRSFSLMCGIVVMVFLPVLKQBYHLVHSEITTHCHDVVDACESPSSFRFY 277
DB 143 SIHKTRCAVACAVVMIISLVAVIPMTFLTSTN--RTNRSACLDITSSDEL-NTIKWYN 199
QY 278 FVSLAFPGFLIPFVIIIFCVTTLIHK-----LKSMDRIWLGVIKAVLLILVIFTI 327
DB 200 LI-LTATTFCLPIVITLCVTTIITHTLHGLQDSCLEKQKAR-----RUTILLALFYV 252
QY 328 CFAPTNILVIHANYHYHN-----DSLYFMYLIALCLGSLNSCLDPFLFYVMS 377
DB 253 CFLPFLHILRVIRIESRLLSISCSIEHQIHEAYIVSRPLAALNTFGNLLYVVS 306
RESULT 40
US-10-270-144-2
; Sequence 2, Application US/10270144
; Publication No. US20030049790A1
```



```
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000750CON
; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/205,196
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
; US-10-270-144-2

Query Match      17.3%; Score 369.5; DB 9; Length 337;
Best Local Similarity 29.6%; Pred. No. 2.2e-24;
Matches 87; Conservative 58; Mismatches 108; Indels 41; Gaps 8;

Qy 107 LSTQVIPAIYLLFVGVGPNIVTL-----WKLRLTKSISLVIFHTNLAIADLLF 157
Db 31 LKMYLPVIYGIIFLVGPGNAVVISYIFKMRPWK-----SSTIIMLNACTDLY 82

Qy 158 CVTLFPFKIAYHLNGNNVFGVMCRITTVVFGNMYCAILITCMGINRYLATAHPETYQ 217
Db 83 LTSLPFLIHYYAGSENIWFGDFMCKFIRFSFHNLYSSILFLCFSIFRYCVIHPMSCF 142

Qy 218 KLPKRSFLLMCGIWMVWVFLYMLPFVILKQEVHLVHSEITTCDDVVVDACESPSRFFYY 277
Db 143 SIHKTRCAVVACAVVWIIISLVAVIPMTFLITSTN--RTNRSACLDLTSSDEL-NTIKWYN 199

Qy 278 FVSLAFPGFLIPFVIIIFCYTTLHK-----LKSXDRIWLGVIKAVLLILVIFTI 327
Db 200 LI-LTATTFCPLVIVTLCTYTIITHTLTHGLQDSCLEKQKAR-----RTILLLLAFYV 252

Qy 328 CFAPTNIILVIHANYYYHNT-----DSLRYMYLIACLGINSCLDPFLYFVMS 377
Db 253 CFLPFLILRVIRIESRLISCSIEHQIHEAYIVSRPLAALNTFGNLLLYVWVS 306
```

Search completed: June 24, 2003, 12:18:35
Job time : 35.4601 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:53:16 ; Search time 14.4822 Seconds
(without alignments)
1165.628 Million cell updates/sec

Title: US-09-208-629F-3

Perfect score: 2136

Sequence: 1 TLYTXQHPVAGSDIKMKIL.....AMARPLXPRRDIEDIHW 407

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1940	90.8	369	1 PAR3_MOUSE	O08675 mus musculu
2	1845.5	77.0	368	1 PAR3_RAT	Q920e1 rattus norv
3	1425.5	66.7	374	1 PAR3_HUMAN	O00254 homo sapien
4	582.5	27.3	420	1 PAR1_XENLA	P47749 xenopus lae
5	581.5	27.2	399	1 PAR2_MOUSE	P55086 mus musculu
6	569	26.6	397	1 PAR2_HUMAN	P55085 homo sapien
7	564.5	26.4	397	1 PAR2_RAT	Q63845 rattus norv
8	540.5	25.3	430	1 PAR1_MOUSE	P30558 mus musculu
9	522.5	24.5	428	1 PAR1_CRILLO	O00991 cricetus
10	512.5	24.0	432	1 PAR1_RAT	P26824 rattus norv
11	508	23.8	425	1 PAR1_PAPHA	P56488 papio hamad
12	507.5	23.8	395	1 PAR4_HUMAN	Q9610 homo sapien
13	501	23.5	435	1 PAR1_HUMAN	P25116 homo sapien
14	486	23.2	395	1 PAR4_RAT	Q920e0 rattus norv
15	494	23.1	396	1 PAR4_MOUSE	O88634 mus musculu
16	431.5	20.2	361	1 EBI2_HUMAN	P32249 homo sapien
17	399	18.7	362	1 P2YR_CHICK	P14996 gallus gall
18	398.5	18.7	367	1 GP17_HUMAN	Q13304 homo sapien
19	395	18.5	345	1 CLT2_PIG	Q95803 sus scrofa
20	395	18.5	362	1 P2YR_MELGA	P49652 meleagris g
21	393	18.4	537	1 P2YR_XENLA	P79928 xenopus lae
22	387.5	18.1	370	1 P2Y9_HUMAN	Q96677 homo sapien
23	383	17.9	308	1 P2Y5_CHICK	P32250 gallus gall
24	382.5	17.9	309	1 CLT2_RAT	Q924t9 rattus norv
25	378.5	17.7	346	1 CLT2_HUMAN	Q9875 homo sapien
26	378.5	17.7	373	1 P2Y2_MOUSE	P35833 mus musculu
27	378	17.7	337	1 CLT1_HUMAN	Q9Y271 homo sapien
28	377.5	17.7	361	1 P2Y4_RAT	Q35811 rattus norv
29	377.5	17.7	373	1 P2YR_RAT	P49651 rattus norv
30	376.5	17.6	373	1 P2YR_HUMAN	P47900 homo sapien
31	372.5	17.4	361	1 P2Y4_MOUSE	Q9JJ57 mus musculu
32	372	17.4	373	1 P2YR_BOVIN	P48042 bos taurus
33	367	17.2	381	1 GP34_HUMAN	Q9upc5 homo sapien

RESULT 1

PAR3_MOUSE

ID PAR3_MOUSE STANDARD; PRT; 369 AA.

AC O08675;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-

DE like 2) (Coagulation factor II receptor-like 2).

GN P2RL2 OR PAR3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6;

RX MEDLINE=97242411; PubMed=9087410;

RA Ishihara H., Connolly A.J., Zeng D., Kahn M.L., Zheng Y.-W.,

Timmons C., Tram T., Coughlin S.R.;

RT "Protease-activated receptor 3 is a second thrombin receptor in

RT humans.";

RL Nature 386:502-506 (1997).

CC -!- FUNCTION: High affinity receptor for activated thrombin coupled to

CC G proteins that stimulate phosphoinositide hydrolysis. May play a

CC role in platelets activation.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- PTM: A proteolytic cleavage generates a new amino terminus that

CC functions as a tethered ligand.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U92972; AAC53137.1; --

DR HSSP; P34996; 1DDD.

DR MGD; MGI:1298208; P2rl2.

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7tm1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G-PROTEIN RECF_F1_1; FALSE_NEG.

DR PROSITE; PS50362; G-PROTEIN_RECF_F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;

KW Blood coagulation.

FT SIGNAL 1 19 POTENTIAL.

FT PROPEP 20 37 REMOVED FOR RECEPTOR ACTIVATION (BY

FT SIMILARITY).

FT CHAIN 38 369 PROTEINASE ACTIVATED RECEPTOR 3.

FT DOMAIN 38 93 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 94 119 1 (POTENTIAL).

FT DOMAIN 120 127 CYTOPLASMIC (POTENTIAL).

ALIGNMENTS

```

FT TRANSMEM 128 147 2 (POTENTIAL).
FT DOMAIN 148 166 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 167 188 3 (POTENTIAL).
FT DOMAIN 189 205 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 206 229 4 (POTENTIAL).
FT DOMAIN 230 258 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 260 279 5 (POTENTIAL).
FT DOMAIN 280 296 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 297 321 6 (POTENTIAL).
FT DOMAIN 322 335 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 336 360 7 (POTENTIAL).
FT DOMAIN 361 369 CYTOPLASMIC (POTENTIAL).
FT SITE 37 38 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT DISULFID 165 244 BY SIMILARITY.
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 369 AA; 41707 MW; F4166BF766D073DB CRC64;

Query Match 90.0%; Score 1940; DB 1; Length 369;
Best Local Similarity 100.0%; Pred. No. 2.9e-101;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 MKLILVAAGLLPLPVTVCOSGINVSDNSAKPTLTIKSPNGGPONTFFEPFLSDIEGWTG 76
DB 1 MKLILVAAGLLPLPVTVCOSGINVSDNSAKPTLTIKSPNGGPONTFFEPFLSDIEGWTG 60

QY 77 ATTTIKAECPDSISTLHVNNATIGYLRSLSTQVIPAIIYLLFVVGVPNTVTLWKL 136
DB 61 ATTTIKAECPDSISTLHVNNATIGYLRSLSTQVIPAIIYLLFVVGVPNTVTLWKL 120

QY 137 RTKSISLVIFHTNLAIDLFCVTLFPKIAHYHLNGNNVFGVWCMRTTTFVFGNMYCAI 196
DB 121 RTKSISLVIFHTNLAIDLFCVTLFPKIAHYHLNGNNVFGVWCMRTTTFVFGNMYCAI 180

QY 197 LILTCGGINRYLATAPFTYQKLPKRSFSLMCGIWMVFLYMLPFVILKQEHYLVHSE 256
DB 181 LILTCGGINRYLATAPFTYQKLPKRSFSLMCGIWMVFLYMLPFVILKQEHYLVHSE 240

QY 257 ITTCHDVVDACESPSSFRFYVSLAFPGFLIPFVILIIICVYTLHLKSKDRIMWLYTK 316
DB 241 ITTCHDVVDACESPSSFRFYVSLAFPGFLIPFVILIIICVYTLHLKSKDRIMWLYTK 300

QY 317 AVLLILVIFTCFAPTNIILVIHANYHYNTDSLYFMVLIACLGSLNSCLDPPFLYFVM 376
DB 301 AVLLILVIFTCFAPTNIILVIHANYHYNTDSLYFMVLIACLGSLNSCLDPPFLYFVM 360

QY 377 SKYVDQLNP 385
DB 361 SKYVDQLNP 369

```

RESULT 2

```

PAR3_RAT ID PAR3_RAT STANDARD; PRT; 368 AA.
AC Q920E1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-like 2) (Coagulation factor II receptor-like 2).
GN F2RL2 OR PAR3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Chien E.K., Marietti S., Mendoza J., Phillippe M.;
RT "Cloning of the rat protease activated receptor isoforms 3 and 4.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for activated thrombin coupled to G proteins that stimulate phosphoinositide hydrolysis.

```

```

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: A proteolytic cleavage generates a new amino terminus that functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; AF310076; AAL36789.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PROSITE; PS00237; G-PROTEIN RECF_F1_1; FALSE_NEG.
CC PROSITE; PS00262; G-PROTEIN RECF_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Blood coagulation.
CC SIGNAL 1 21 POTENTIAL.
CC PROPEP 22 37 REMOVED FOR RECEPTOR ACTIVATION (BY SIMILARITY).
CC CHAIN 38 368 PROTEINASE ACTIVATED RECEPTOR 3.
CC DOMAIN 38 93 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 94 114 1 (POTENTIAL).
CC DOMAIN 115 123 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 124 144 2 (POTENTIAL).
CC DOMAIN 145 166 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 167 187 3 (POTENTIAL).
CC DOMAIN 188 208 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 209 229 4 (POTENTIAL).
CC DOMAIN 230 257 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 258 278 5 (POTENTIAL).
CC DOMAIN 279 300 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 301 321 6 (POTENTIAL).
CC DOMAIN 322 338 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 339 359 7 (POTENTIAL).
CC DOMAIN 360 368 CYTOPLASMIC (POTENTIAL).
CC SITE 37 38 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
CC DISULFID 164 243 BY SIMILARITY.
CC CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 368 AA; 41795 MW; CE0E94EDA3B80EF1 CRC64;

Query Match 77.0%; Score 1645.5; DB 1; Length 368;
Best Local Similarity 83.5%; Pred. No. 6e-85;
Matches 309; Conservative 31; Mismatches 25; Indels 5; Gaps 3;

QY 15 IKMKILILVAAGLLPLPVTVCOSGI-NVSDNSAKPTLTIKSPNGGPONTFFEPFLSDIEG 73
DB 1 MEMKVILLVGVRLLPLPVTVCOSGMKHSVDSNA---LTAESFNGN-EHSFEFFPLSDIEG 56

QY 74 WTGATTTIKAECPDSISTLHVNNATIGYLRSLSTQVIPAIIYLLFVVGVPNTVTLWK 133
DB 57 WTGATTTIKAECPDSISTLHVNNATIGYLRSLSTQVIPAIIYLLFVVGVPNTVTLWK 116

QY 134 LSLRTKSLVIFHTNLAIDLFCVTLFPKIAHYHLNGNNVFGVWCMRTTTFVFGNMY 193
DB 117 LSSRTKSLVIFHTNLAIDLFCVTLFPKIAHYHLNGNDVFGVWCMRTTTFVFGNMY 176

QY 194 CAILILTCGGINRYLATAPFTYQKLPKRSFSLMCGIWMVFLYMLPFVILKQEHYLV 253
DB 177 CAILILTCGGINRYLATAPFTYQKLPKRNFTLLMCGVWMVFLYMLPLAILKQEHYLV 236

QY 254 HSEITTCCHDVVDACESPSSFRFYVSLAFPGFLIPFVILIIICVYTLHLKSKDRIMW 313
DB 237 QGITTTCCHDVVDACESPPLFPQFYVSLAFPGFLIPFVSVVFCYTTLHLKNAQKKWL 296

QY 314 YIKAVLLILVIFTCFAPTNIILVIHANYHYNTDSLYFMVLIACLGSLNSCLDPPFLY 373
DB 297 YIKAVLLILVIFTCFAPTNIILVIHANYHYNTDSLYFMVLIACLGSLNSCLDPPFLY 356

QY 374 FYMSKVVDQL 389

```

Db	357 FMSKIVDQL 366	:	
RESULT 3			
PAR3 HUMAN			
ID	PAR3_HUMAN	STANDARD;	PRT; 374 AA.
AC	O00254;		
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-like 2) (Coagulation factor II receptor-like 2).		
DE	F2RL2 OR PAR3.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	(1)		
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF THR-39 AND PHE-40.		
RX	MEDLINE=97242411; PubMed=9087410;		
RA	Ishihara H., Connolly A.J., Zeng D., Kahn M.L., Zheng Y.-W.,		
RA	Timmons C., Tram T., Coughlin S.R.;		
RT	"Protease-activated receptor 3 is a second thrombin receptor in humans.";		
RT	Nature 386:502-506(1997).		
RL	(2)		
RN	SEQUENCE FROM N.A., AND VARIANTS SER-15; VAL-177 AND ASP-250.		
RP	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Yi Q.,		
RA	Nickerson D.A.;		
RA	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.		
RL	(3)		
RN	TISSUE SPECIFICITY.		
RP	MEDLINE=98279023; PubMed=9614115;		
RX	Schmidt V.A., Nieman W.C., Maglott D.R., Cupit L.D., Moskowitz K.A.,		
RA	Walner J.A., Bahou W.F.;		
RA	"The human proteinase-activated receptor-3 (PAR-3) gene.		
RT	Identification within a PAR gene cluster and characterization in		
RT	vascular endothelial cells and platelets.";		
RL	J. Biol. Chem. 273:15061-15068(1998).		
RN	(4)		
RP	FUNCTION.		
RX	MEDLINE=99178892; PubMed=10079109;		
RA	Kahn M.L., Nakanishi-Matsui M., Shapiro M.J., Ishihara H.,		
RA	Coughlin S.R.;		
RT	"Protease-activated receptors 1 and 4 mediate activation of human		
RT	platelets by thrombin.";		
RL	J. Clin. Invest. 103:879-887(1999).		
CC	-!- FUNCTION: Receptor for activated thrombin coupled to G proteins		
CC	that stimulate phosphoinositide hydrolysis.		
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-!- TISSUE SPECIFICITY: Highest expression in the megakaryocytes of		
CC	the bone marrow, lower in mature megakaryocytes, in platelets and		
CC	in a variety of other tissues such as heart and gut.		
CC	-!- PTM: A proteolytic cleavage generates a new amino terminus that		
CC	functions as a tethered ligand.		
CC	-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; U92971; AAC51218.1; -		
DR	EMBL; AF374726; AAK51564.1; -		
DR	HSSP; P34996; 1DDD.		
DR	Genew; HGNC:3539; F2RL2.		
DR	MIM; 601919; -		
DR	InterPro; IPR000276; GPCR_Rhodpan.		
DR	Pfam; PF00001; 7tm_1; 1.		

PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.

DR PROSITE; PS00262; G-PROTEIN RECP_F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;

KW Blood coagulation; Polymorphism.

FT SIGNAL 1 21 POTENTIAL.

FT PROPEP 22 38 REMOVED FOR RECEPTOR ACTIVATION (BY

FT SIMILARITY).

FT CHAIN 39 374 PROTEINASE ACTIVATED RECEPTOR 3.

FT DOMAIN 39 94 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 95 120 1 (POTENTIAL).

FT DOMAIN 121 128 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 129 148 2 (POTENTIAL).

FT DOMAIN 149 167 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 168 189 3 (POTENTIAL).

FT DOMAIN 190 206 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 207 230 4 (POTENTIAL).

FT DOMAIN 231 259 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 261 280 5 (POTENTIAL).

FT DOMAIN 281 297 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 298 322 6 (POTENTIAL).

FT DOMAIN 323 336 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 337 361 7 (POTENTIAL).

FT DOMAIN 362 374 CYTOPLASMIC (POTENTIAL).

FT SITE 38 39 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).

FT DISULFID 166 245 BY SIMILARITY.

FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARIANT 15 15 L -> S.

FT /FTID=VAR_012849.

FT M -> V.

FT /FTID=VAR_012850.

FT N -> D.

FT /FTID=VAR_012851.

FT T->P; NO PROTEOLYTIC CLEAVAGE (BY

FT THROMBIN).

FT F->A; ALTERED SIGNAL UPON THROMBIN

FT CLEAVAGE.

FT C45C15A695DDIABB CRC64;

SQ SEQUENCE 374 AA; 42508 MW; 42508 MW; 42508 MW;

Query Match 66.7%; Score 1425.5; DB 1; Length 374;

Best Local Similarity 71.6%; Pred. No. 9.4e-73;

Matches 260; Conservative 44; Mismatches 58; Indels 1; Gaps 1;

QY 17 MKLILVAAGLLFLPVTVCQSGI-NVSDNSAKPTLTIKSENGPONTPEERPLSDIEGWT 75

DB 1 MKALIFAAGLLLLLPFCQSGMENDTNLAKPTLPIKTRGAPNSFEFPFSALEGT 60

QY 76 GATTTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAITYILLFVGVSPSIVTLWKLS 135

DB 61 GATITVKIKPEESASHLVKNATNGYLTSSLTKLIPAITYLLFVGVSPANAVTLWMLF 120

QY 136 LRTKISILVFIHTNIAIADLLFCVTLPKIAYHLNGNNVFGVEVMCRITTVVYGNMYCA 195

DB 121 FRTRSICITVFYTNIAIADFLFCVTLPKIAYHLNGNNVFGVEVLCRATTVIFYGNMYS 180

QY 196 ILILTCMGINRYLATAHPFTYOKLPKRSFSLMCCGIVMMVFLYMLPFVILKQEVHLVHS 255

DB 181 ILLACISINRYLAIVHPFTYRGLPKRYVALVTCGLWATVFLYMLPFILKQEYLVOP 240

QY 256 EITTCDDVVVDACESPSPFRFYFVSLAFPGFLIPVIFCYTTTILHLKSKDRILWGYI 315

DB 241 DITTCDDVHNTCESSPQLYFYFISLAFPGFLIPVILYCYAAIIRTLNAYDHRWLWV 300

QY 316 KAVLILVIFTCFAPTNILVIHANYHNTDSLYFWYLIALCLGSLNSCLDPFLYFV 375

DB 301 KASLLILVIFTCFAPSNILIIHHANYHNTDGLYFYLYIALCLGSLNSCLDPFLYF 360

QY 376 MSK 378

DB 361 MSK 363

```

RESULT 4
PARI_XENLA
ID PARI_XENLA STANDARD; PRT; 420 AA.
AC F47749;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94195429; PubMed=8145852;
RA Gerszten R.E., Chen J., Ishii M., Ishii K., Nanavicz T.,
RA Turk C.W., Vu T.-K.H., Coughlin S.R.;
RT "Specificity of the thrombin receptor for agonist peptide is defined
RT by its extracellular surface.";
RL Nature 368:648-651(1994).
CC -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC G proteins that stimulate phosphoinositide hydrolysis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U09632; AAA18498.1; -
CC InterPro; IPR000276; GPCR_Rhodopsin.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODOPS.
CC PROSITE; PS00237; G-PROTEIN RECP F1_1; 1.
CC PROSITE; PS0262; G-PROTEIN RECP F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Blood coagulation.
CC SIGNAL 1 20 POTENTIAL.
CC PROPEP 21 42 REMOVED FOR RECEPTOR ACTIVATION.
CC CHAIN 43 420 PROTEINASE ACTIVATED RECEPTOR 1.
CC DOMAIN 43 101 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 102 127 1 (POTENTIAL).
CC DOMAIN 128 136 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 137 156 2 (POTENTIAL).
CC DOMAIN 157 175 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 176 197 3 (POTENTIAL).
CC DOMAIN 198 217 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 218 238 4 (POTENTIAL).
CC DOMAIN 239 267 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 268 287 5 (POTENTIAL).
CC DOMAIN 288 310 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 311 333 6 (POTENTIAL).
CC DOMAIN 334 345 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 346 369 7 (POTENTIAL).
CC DOMAIN 370 420 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 295 300 POLY-SER.
CC SITE 42 43 CLEAVAGE (BY THROMBIN).
CC DISULFID 174 253 BY SIMILARITY.
CC CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC MUTAGEN 42 42 K-A: LACK OF ACTIVATION BY THROMBIN.
CC SEQUENCE 420 AA; 47435 MW; D5163F6AFE12372 CRC64;
Query Match 27.3%; Score 582.5; DB 1; Length 420;
Best Local Similarity 32.7%; Pred. No. 5.3e-26;

```

Matches 124; Conservative 82; Mismatches 138; Indels 35; Gaps 11;

QY 19 ILILVAAGLFLPVTVCOSGINVDSNAK-----PLTIKSFN--GGPONTTEERPLSDIE 72
 DB 7 LLLLLLLTLLGAMGSLCLAN---SDTQAGAHNNMTIKTRIPDSDSEFEIIPWDEL 63
 QY 73 GWTGATTIKAECPEDSISTL--HVNNTATIGYLRSSLSSTQVIPAIIYILLFVVGVPNS-- 127
 DB 64 E-SGEGSDQAPVRSARKPIRRNITKEAEQYLSQWLTKFVPSLYTVVIVGLPLNLLA 122
 QY 128 -IVTLWKLSTKTSISLVFHTNIAIADLLFCVTLFPKIAHYHLNNGNWFVGEVMCRITTV 186
 DB 123 IIFLFKMKVRKPA---VYVMLNLAIDVFFVSLVFPFKIAYHLSGNDWLFGEVGMCRIVTA 179
 QY 187 VFYGNMYCAILLITCMGINRYLATAPFTYQKLPKRSFSLMCGIVVMVFLYMLPFVIL 246
 DB 180 IFYCNMYCSVLLIASISVDRLAVVYPMHSLSWRMSRAYMACSFIMLISIASTIPLIVT 239
 QY 247 KOEYHLVHSEITTHDVEDDACESPSFRFYFVSLAFFGFLIPFVIIIFCYTTLHKLS 306
 DB 240 EQTKIPRLDITTHCDVLD-LKDLKDFVYVYFSSFCLLFFVFPFIITTCYIGIIRSLSS 298
 QY 307 -----KDRILGYIKAVLLILVIFTCFAPTNIIIVHHANYHNTDSLYFWYLIA 358
 DB 299 SSIENSKCKTRA---LFLAVVLCVFCICGPTNVLFTH---YLOEANEFLYFAYILS 351
 QY 359 LCLGSLNSCLDPFLYFVMS 377
 DB 352 ACVGSVSCCLDPLIYYAS 370

RESULT 5
 PARI_MOUSE
 ID PARI_MOUSE STANDARD; PRT; 399 AA.
 AC P55086;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 2 precursor (PAR-2) (Thrombin receptor-
 DE like 1) (Coagulation factor II receptor-like 1).
 GN F2RL1 OR PAR2 OR GPR11 OR GPCR11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95197620; PubMed=7890726;
 RA Nystedt S., Larsson A.-K., Aaberg H., Sundelin J.;
 RT "The mouse proteinase-activated receptor-2 cDNA and gene. Molecular
 RT cloning and functional expression";
 RL J. Biol. Chem. 270:5950-5955(1995).
 CC -!- FUNCTION: Receptor for trypsin and trypsin-like enzymes coupled to
 CC G proteins that stimulate phosphoinositide hydrolysis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
 CC functions as a tethered ligand.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z48043; CAA88097.1; -
 CC HSSP; P34996; 1DDD.
 CC MGI; MGI:101910; F2rl1.
 CC InterPro; IPR000276; GPCR_Rhodopsin.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCRHHODOPS.

DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; FALSE_NEG.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT PROPEP 26 38 REMOVED FOR RECEPTOR ACTIVATION
 FT CHAIN 39 399 (BY SIMILARITY).
 FT DOMAIN 39 77 PROTEINASE ACTIVATED RECEPTOR 2.
 FT TRANSSEM 78 103 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 104 112 1 (POTENTIAL).
 FT TRANSSEM 113 132 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 133 151 2 (POTENTIAL).
 FT TRANSSEM 152 173 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 174 192 3 (POTENTIAL).
 FT TRANSSEM 193 213 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 214 243 4 (POTENTIAL).
 FT TRANSSEM 244 262 5 (POTENTIAL).
 FT DOMAIN 263 287 6 (POTENTIAL).
 FT TRANSSEM 288 310 7 (POTENTIAL).
 FT DOMAIN 311 325 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 326 349 7 (POTENTIAL).
 FT DOMAIN 350 399 CYTOPLASMIC (POTENTIAL).
 FT SITE 385 394 POLY-SER.
 FT DISULFID 38 39 CLEAVAGE (BY TRYPSIN) (POTENTIAL).
 FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 399 AA; 44752 MW; A93749425ED0B194 CRC64;

Query Match 27.2%; Score 581.5; DB 1; Length 399;
 Best Local Similarity 37.6%; Pred. No. 5.8e-26;
 Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY 95 VNNATIGYLRSSLSQVIAIVLLFVGVGPNIVTLWKLRLT-KSISLVIFHTNLATA 153
 DB 63 IDFSASITGKLTIVFLPVPVIVIVGIPGNSGMALFPLRTKKHPAVYMANLAA 122
 QY 154 DLLFCVTLPLFKIAYHLGNVNVFGEVYMCRTTIVFYGNMYCAIILITCMGINRYLATAP 213
 DB 123 DLLSVIWFPLKISYHLGNVNVVGEALCKVLGFFYGNMYCSILFMTCLSVQRYVYVNP 182
 QY 214 FTYQKLPKPSFSLMCGIWMVFLYMLPFLVKQYHLVHSEITTHDVPDACESPSF 273
 DB 183 MGHPR-KKANIAGVSLAIWLFILFTIPLYMKQTIYIPAKNITTHDVLPE-EVLVGD 240
 QY 274 RYFYVSLAFEGFLIEPVIIFCYTTLI-----HKLKSKRIMGLGYKAVLLIIV 323
 DB 241 MNYFLSLAIGVFLPALUTASAYLVMIKTLRSSAMDEHSEKKRQRA----IRLIITVLA 296
 QY 324 IFTICFAPTNILVVIHANYYYHNTDSLXYFMYLIALCLGSLNSCLDPFLYFVMSK 378
 DB 297 MYFICFAPSNLLLVHVFLEIKTQROSHVYALYVALCLTSLNSCIDPFVYFVSK 351

RESULT 6
 ID PAR2 HUMAN STANDARD; PRT; 397 AA.
 AC P55085; Q13317; Q13346;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 2 precursor (PAR-2) (Thrombin receptor-like 1) (Coagulation factor II receptor-like 1).
 GN P2RL1 OR PAR2 OR GPR11.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96048032; PubMed=7556175;
 RA Nystedt S., Emlissson K., Larsson A.-K., Stroembeck B., Sundelin J.;
 RT "Molecular cloning and functional expression of the gene encoding the

human proteinase-activated receptor 2.";
 Eur. J. Biochem. 232:84-89(1995).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=9617879; PubMed=8615752;
 RA Boehm S.K., Kong W., Broemme D., Smeekens S.P., Anderson D.C.,
 RA Connolly A.J., Kahn M.L., Nelken N.A., Coughlin S.R., Payan D.G.,
 RA Bunnett N.W.;
 RT "Molecular cloning, expression and potential functions of the human
 proteinase-activated receptor-2.";
 Biochem. J. 314:1009-1016(1996).
 [3]
 RP SEQUENCE OF 29-397 FROM N.A.
 RX MEDLINE=96379236; PubMed=8784787;
 RA Kahn M.L., Ishii K., Kuo W.L., Piper M., Connolly A.J., Shi Y.P.,
 RA Wu R., Lin C.C., Coughlin S.R.;
 RT "Conserved structure and adjacent location of the thrombin receptor
 and protease-activated receptor 2 genes define a protease-activated
 receptor gene cluster.";
 Mol. Med. 2:349-357(1996).
 [4]
 RP SEQUENCE FROM N.A., AND VARIANTS PHE-21; GLN-270 AND ALA-291.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary, and Pancreas;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Receptor for trypsin and trypsin-like enzymes coupled to
 G proteins that stimulate phosphoinositide hydrolysis. May have a
 role in the regulation of vascular tone.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN TISSUES WITH ESPECIALLY
 HIGH LEVELS IN PANCREAS, LIVER, KIDNEY, SMALL INTESTINE, AND
 COLON. MODERATE EXPRESSION IS DETECTED IN MANY ORGANS, BUT NONE IN
 BRAIN OR SKELETAL MUSCLE.
 CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
 functions as a tethered ligand.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 or send an email to license@isb-sib.ch).

 EMBL; Z49993; CAA90290.1; -;
 EMBL; Z49994; CAA90290.1; JOINED.
 EMBL; U34038; AAB4781.1; -;
 EMBL; U36753; AAA90957.1; -;
 EMBL; AF400075; AAK77914.1; -;
 EMBL; BC012453; AAH12453.1; -;
 EMBL; BC018130; AAH18130.1; -;
 HSSP; P34996; 1DDD.
 DR Genew; HGNC:3538; P2RL1.
 DR MIM; 600933; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHOOPS
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; FALSE_NEG.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Polymorphism.
 FT SIGNAL 1 25 POTENTIAL.
 FT PROPEP 26 36 REMOVED FOR RECEPTOR ACTIVATION
 FT CHAIN 37 397 (BY SIMILARITY).
 FT DOMAIN 37 75 PROTEINASE ACTIVATED RECEPTOR 2.
 FT EXTRACELLULAR (POTENTIAL).


```
QY 214 FTYOKLPKRSFSLMCGIWMVFLVLPVILKQBYHLVHSEITTHCHVDVACESPSSF 273
Db 181 MHSR-KRANIAVGSALWLAIFLTVIPLYVMRQTIYPALNITTHCHVLP-EVLVDG 238
QY 274 RFYVFSVLAFFGLIFPFIIFCYTTLIHKLK-----KDRILWGYIKAVLLILVIFI 327
Db 239 MESYFSLAIGVFLPALITASNYLMIKTLRSSAMDEHSEKKRRRAIRLITVLSMYFI 298
QY 328 CFAPNTNIIIVHANYHTNDSLYPMYLIACLSLNSCLDPFLYFVMSK 378
Db 299 CFAPSNVLLVWHYFLIKSQSHVYALVALCISLNSCIDPFVYFVYSK 349

RESULT 8
PARI_MOUSE
ID PARI_MOUSE STANDARD; PRT; 430 AA.
AC P3058; P97507;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
GN F2R OR PARI OR CF2R.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Coughlin S.R.;
RL Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=96379236; PubMed=8784787;
RA Kahn M.L., Ishii K., Kuo W.L., Piper M., Connolly A., Shi Y.P.,
RA Wu R., Lin C.C., Coughlin S.R.;
RT "Conserved structure and adjacent location of the thrombin receptor
RT and protease-activated receptor 2 genes define a protease-activated
RT receptor gene cluster."
RL Mol. Med. 2:349-357(1996).
RN [3]
RP SEQUENCE OF 1-74 FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=96327649; PubMed=8678993;
RA Xue J., Jenkins N.A., Gilbert D.J., Copeland N.G., Sadler J.E.;
RT "Structure and localization of the thrombin receptor gene on mouse
RT chromosome 13."
RL Mamm. Genome 7:625-626(1996).
CC -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC G proteins that stimulate phosphoinositide hydrolysis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; L03529; AAA040438.1;
CC EMBL; U36757; AAB38308.1;
CC EMBL; U36756; AAB38308.1; JOINED.
CC EMBL; U55076; AAB00198.1;
CC EMBL; U55075; AAB00198.1; JOINED.
CC HSSP; P34996; 1DDD.
CC MGD; MGI:101802; F2r.
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm_1; 1.
```

```
DR PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.
DR PROSITE; PS50262; G-PROTEIN RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Blood coagulation.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 41 REMOVED FOR RECEPTOR ACTIVATION (BY
FT SIMILARITY).
FT CHAIN 42 430 PROTEINASE ACTIVATED RECEPTOR 1.
FT DOMAIN 42 107 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 108 133 1 (POTENTIAL).
FT DOMAIN 134 142 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 143 162 2 (POTENTIAL).
FT DOMAIN 163 181 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 182 203 3 (POTENTIAL).
FT DOMAIN 204 223 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 224 244 4 (POTENTIAL).
FT DOMAIN 245 273 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 274 293 5 (POTENTIAL).
FT DOMAIN 294 316 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 317 339 6 (POTENTIAL).
FT DOMAIN 340 354 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 355 379 7 (POTENTIAL).
FT DOMAIN 380 430 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 60 65 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 88 91 POLY-PRO.
FT SITE 41 42 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT DISULFID 180 259 BY SIMILARITY.
FT CARBOHYD 67 67 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 80 80 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 162 162 F -> S (IN REF. 1).
FT CONFLICT 189 189 G -> Y (IN REF. 1).
FT CONFLICT 223 223 R -> G (IN REF. 1).
FT CONFLICT 262 262 V -> L (IN REF. 1).
FT CONFLICT 365 365 S -> T (IN REF. 1).
SQ SEQUENCE 430 AA; 47790 MW; 395FD64FAE52C9BF CRC64;
Query Match 25.3%; Score 540.5; DB 1; Length 430;
Best Local Similarity 31.6%; Pred. No. 1.1e-23;
Matches 124; Conservative 79; Mismatches 137; Indels 53; Gaps 13;
QY 20 LILVAAG-----LLFLPVTVCSQGINVSDNSAKTLTIKSF--NGGPONTTEEPPLSDI 71
Db 6 LLILVALGLSLGCLLSRVPMSPQESERTATNP-----RSFFLRNPSENTFLVPLGDE 61
QY 72 EG-----WTGATTTIKAECPDISITLHVNNATIGVLRSSLSSTQVIPAIVILLFVVG 123
Db 62 EEEKNESVLLLEGRAVYLNISLPHTPPPPPISEASGLTSPWLTLPMPSTVITFVVS 121
QY 124 VPSNTVTLWKLRLTK-SISLVIFHTNLAIAADLLFCVTLPPKAIYHLGNMNVGEMVCR 182
Db 122 LPLNLVLAIAVFLRMKKKPAVVYMLHAMADVLSVLPFKISYVFSCTDMQFGSGMCR 181
QY 183 ITTVFYGNMYCAIILITCMGINRVLAHP---FTYOKLPKRSFSLMCGIWMVFLY 239
Db 182 FATAFYGNMYASIMLMTVISIDREFLAVVYPTQSLSWRTLGRANFT---CVVIWMAIMG 238
QY 240 MLPFVLKQYHLVHSEITTHCHVDVACESPSPRFYFVSLAFFGLIFPFIIFCYTT 299
Db 239 VVPLLEKQSTRVPGNLITTHCHDVUSE-NLMQGFYSYFSAFSAFFLPLVLSVTCYTS 297
QY 300 LIHKLS-----KDRILWGYIKAVLL---ILVIFTICFAPNTNIIIVHANY--- 345
Db 298 IIRCLSSAVANRSKSR-----ALFLSAAVFCIFVCGFTNVLIVH---YLFLSD 347
QY 346 -HNTDSLYPMYLIACLSLNSCLDPFLYFVMS 377
Db 348 SPGETAAYFAVLLCVCSVSCCIDPLIYYVAS 380

RESULT 9
PARI_CRIO
ID PARI_CRIO STANDARD; PRT; 428 AA.
```

Q00991; Q60461;
 01-APR-1993 (Rel. 25, Created)
 01-NOV-1997 (Rel. 35, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
 GN F2R OR PAR1.
 OS Cricetulus longicaudatus (Long-tailed hamster) (Chinese hamster).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetulus.
 NCBI_TaxId=10030;
 [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=91348247; PubMed=1652467;
 RA Rasmussen U.B., Vouret-Craviari V., Jallat S., Schlesinger Y.,
 RA Pages G., Pavirani A., Lecocq J.P., Pouyssegur J.,
 RA Oberghen-Schilling E.;
 RT "cDNA cloning and expression of a hamster alpha-thrombin receptor
 RT coupled to Ca2+ mobilization.";
 RL FEBS Lett. 288:123-128(1991).
 RN [2]
 RP SEQUENCE OF 42-428 FROM N.A.
 RX MEDLINE=96028007; PubMed=7488069;
 RA Hartmann T., Grace M.B., Buzard G.S., Ruoss S.J.;
 RA "Thrombin receptor polymorphism in Chinese hamster";
 RL Biochem. Biophys. Res. Commun. 215:974-980(1995).
 CC -1- FUNCTION: High affinity receptor for activated thrombin coupled to
 CC G proteins that stimulate phosphoinositide hydrolysis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PTM: A proteolytic cleavage generates a new amino terminus that
 CC functions as a tethered ligand.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X61958; CAA43957.1; -;
 CC EMBL; U34047; AAB6747.1; -;
 CC PIR; S17148; S17148.
 CC HSP; P34996; 1DDD.
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 CC PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 CC Blood coagulation; Polymorphism.
 KW SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 41 REMOVED FOR RECEPTOR ACTIVATION (BY
 FT SIMILARITY).
 FT CHAIN 42 428
 FT DOMAIN 42 105
 FT TRANSMEM 106 131
 FT DOMAIN 132 140
 FT TRANSMEM 141 160
 FT DOMAIN 161 179
 FT TRANSMEM 180 201
 FT DOMAIN 202 221
 FT TRANSMEM 222 242
 FT DOMAIN 243 271
 FT TRANSMEM 272 291
 FT DOMAIN 292 314
 FT TRANSMEM 315 337
 FT DOMAIN 338 352
 FT TRANSMEM 353 377
 FT DOMAIN 378 428
 FT DOMAIN 60 63 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 301 304 POLY-SER.

FT SITE 41 42 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT DISULFID 178 257 BY SIMILARITY.
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 84 84 MISSING (IN AN ALLELE).
 FT CONFLICT 384 384 H -> T (IN REF. 1).
 SQ SEQUENCE 428 AA; 47602 MW; E5492AB2CD877E2F CRC64;
 Query Match 24.5%; Score 522.5; DB 1; Length 428;
 Best Local Similarity 32.7%; Pred. No. 1.1e-22;
 Matches 129; Conservative 80; Mismatches 128; Indels 57; Gaps 16;
 QY 20 LILVAAG-----LLFLPVTVCQGINVSDNSAKPTLTIKSF---NGGPQNTFEFFPLSD 70
 DB 6 LLLVAAGLSLGCPLLSRVPVRQPESEMTDATVNP-----RSFFLRNPG-ENTFEIPLGD 60
 QY 71 IEGWTGATT-----TIKACPEPDSISTLHVNNATIGYLRSSLSLSTQVIPAIYILLFVVG 123
 DB 61 EEEKNESTLPESGRAIYLNKSHSPAPLAPFISEDAS-CYLTPSWLRLFIPSYVTFVVS 119
 QY 124 VPSNI--VTLMKLSLRTKISILVIFHTNLAIADLLFCVTLPFKIAYHLNNGNNGVGEVWC 181
 DB 120 LPLNLAIAVFLVKMKVKK-PAVVYMLHLAMADLVLSVLPKISYFSGSDMQFGSGMC 178
 QY 182 RITTVFVGNMYCATILITCGINRYLATAHP---FTYQKLPKRSFSLMCGIWMVWVFL 238
 DB 179 RFATAAFYCNMYASIMLVISIDRFILAVVPIQSLSWRTLGANFT---CLVIMWAIM 235
 QY 239 YMLPEVLKQEHVHSEITTCVDVDACESPSPFRFYFVSLAFPGFLIPFVILFCYT 298
 DB 236 GWPPLLKEQITRVEGLNITTCVDVLE-TLLQGFYSYFSAFSAVFFLPLIISTICVM 294
 QY 299 TLIHLKLS-----KDRIMLYIKAVLL---ILVIFTICFAPNIIILVHHANYHYN 347
 DB 295 SIIRCLSSSVANRSKSR-----ALFLSAAVCFVFCVPGPTNVLIMH---YLLLS 344
 QY 348 ---TDSLYFMYLIALCLGSLNSCLDPLFYFMS 377
 DB 345 DSPATEKAYFAYLLCCVCSVSSCCIDPLIYYAS 378
 RESULT 10
 PAR1_RAT
 ID PAR1_RAT STANDARD; PRT; 432 AA.
 AC P26824;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
 GN F2R OR PAR1.
 OS Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
 NCBI_TaxId=10116;
 [1]
 SEQUENCE FROM N.A.
 RP TISSUE=Aortic smooth muscle;
 RC MEDLINE=92381002; PubMed=1324917;
 RA Zhong C., Hayzer D.J., Corson M.A., Wick K., Runge M.S.;
 RT "Molecular cloning of the rat vascular smooth muscle thrombin
 RT receptor. Evidence for in vitro regulation by basic fibroblast growth
 RT factor.";
 RL J. Biol. Chem. 267:16975-16979(1992).
 CC -1- FUNCTION: High affinity receptor for activated thrombin coupled to
 CC G proteins that stimulate phosphoinositide hydrolysis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PTM: A proteolytic cleavage generates a new amino terminus that
 CC functions as a tethered ligand.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

-----
EMBL; M81642; AAA42274.1; -
DR PIR; A43448; A43448.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOPOSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Blood coagulation.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 45 REMOVED FOR RECEPTOR ACTIVATION (BY
  SIMILARITY).
FT CHAIN 46 432 PROTEINASE ACTIVATED RECEPTOR 1.
FT DOMAIN 46 109 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 110 135 1 (POTENTIAL).
FT DOMAIN 136 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 184 2 (POTENTIAL).
FT DOMAIN 165 183 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 184 205 3 (POTENTIAL).
FT DOMAIN 206 225 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 226 246 4 (POTENTIAL).
FT DOMAIN 247 275 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 276 295 5 (POTENTIAL).
FT DOMAIN 296 318 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 319 341 6 (POTENTIAL).
FT DOMAIN 342 357 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 358 381 7 (POTENTIAL).
FT DOMAIN 382 432 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 64 68 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 87 93 POLY-PRO.
FT SITE 45 46 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT DISULFID 182 261 BY SIMILARITY.
FT CARBOHYD 69 69 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 432 AA; 48280 MW; DD032B97ABA4A605 C64;

Query Match
Best Local Similarity 24.0%; Score 512.5; DB 1; Length 432;
Matches 126; Conservative 75; Mismatches 134; Indels 65; Gaps 13;

QY 20 LILVAGLLF-----LPVTVCSQGINVSDNSAKPTLTIKSFNGGPGONTPEERPLSD- 70
Db 6 LLLVAVGLSLGPLLSSRVPMPQSESRMYATPYATPNRSPFLRNPSEDTEQEPFLGDE 65

QY 71 -----IEGWTGATTTKAECP-----EDSISTLVHNNATIGYLRSSLSQTVP 113
Db 66 EKNESIPLEG--RAVYLNKSRFPMPPPFISED-----SGYLTSPWLTLP 113

QY 114 AIYILLFVGVPSNIWTLKLSLRTK-SISLVIFHTNLAIADLLFCVTLPPFKIAYHLNGN 172
Db 114 SVYTFVFIIVSLPLNLAIAVFRMVKRPAVYMLHLMADVLVFSVLPFKISYVFSGT 173

QY 173 NWVFGVWCRITTVVYGNMYCAILITCGINRYLATAPH---FTYQKLPRKRSFLLMC 229
Db 174 DWQFGGMRCFATAACYCNMYASIMLTVISIDRFLAVVPIQSLSWRTLGRANFT---C 230

QY 230 GIWVWVFLYMLPFVILKQYHLVHSEITTHDWDVACSPSPFRFYVSLAFFGFLP 289
Db 231 VWIWNMAINGVPLLEKEQTTPVGLNITTHDVLNE-TLLHGFSYYSFAESAIFFLVP 289

QY 290 FVLIIFCYTTLHLKLS-----KDRWLGVYKAVLL---ILVIFTICFAPTNIILVI 338
Db 290 LIISTVCYTSIIRCLSSSAVANRKKSR-----ALFLSAAVFCIFVFCGPTNVLIV 342

```

RESULT 11

```

PARI_PAPHA
ID PARI_PAPHA STANDARD; PRT; 425 AA.
AC P56488;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUN-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 1 precursor (PARI-1) (Thrombin receptor).
GN F2R OR PARI OR BTHR12.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OC NCBI_TaxID=9557;
RN [1]
SEQUENCE FROM N.A.
RP Shoji M., Hayzer D.J., Hanson S.R.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RL -!- FUNCTION: High affinity receptor for activated thrombin coupled to
  G proteins that stimulate phosphoinositide hydrolysis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
  functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
```

```

QY 339 HHANYI-YHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
  : : : : : : : : : : : : : : : : : : : : : : : : :
Db 343 HYLILSDSPGTETAYFAYLLCVCVTSVASCIDPLIYYVAS 382
  : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

SIGNAL 1 26 POTENTIAL.
PROPEP 27 41 REMOVED FOR RECEPTOR ACTIVATION.
CHAIN 42 425 PROTEINASE ACTIVATED RECEPTOR 1.
DOMAIN 42 102 EXTRACELLULAR (POTENTIAL).
TRANSMEM 103 128 1 (POTENTIAL).
DOMAIN 129 137 CYTOPLASMIC (POTENTIAL).
TRANSMEM 138 157 2 (POTENTIAL).
DOMAIN 158 176 EXTRACELLULAR (POTENTIAL).
TRANSMEM 177 198 3 (POTENTIAL).
DOMAIN 199 218 CYTOPLASMIC (POTENTIAL).
TRANSMEM 219 239 4 (POTENTIAL).
DOMAIN 240 268 EXTRACELLULAR (POTENTIAL).
TRANSMEM 269 288 5 (POTENTIAL).
DOMAIN 289 311 CYTOPLASMIC (POTENTIAL).
TRANSMEM 312 334 6 (POTENTIAL).
DOMAIN 335 350 EXTRACELLULAR (POTENTIAL).
TRANSMEM 351 374 7 (POTENTIAL).
DOMAIN 375 425 CYTOPLASMIC (POTENTIAL).
SITE 57 60 ASP/GLU-RICH (ACIDIC).
SITE 41 42 CLEAVAGE (BY THROMBIN).
DISULFID 175 254 BY SIMILARITY.
CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 75 75 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 250 250 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).

```


CC -!- FUNCTION: High affinity receptor for activated thrombin coupled to

Db 61 KNEISGLTEYRLVSKNSPLQKLPAPFISDASGYLTSSWLTFLVPSVYTGFWVWSLPLN 120
 QY 128 I--VTMLKLSLRKTSISLVIFHNLAIDLPCVTLPKIAHYLNGNNVFGVMCRIT 185
 Db 121 IMAIVFVLKXVKK--PAVVMUHLATADLVFVSLPFLKISYFSGSDMQFSELCRFT 179
 QY 186 VVFGYNNYCAILLTCMGINRYLATAHP---FTYQKLPKRSFSLMCGIWMVFLYMLP 242
 Db 180 AAFYCNMYASILLMTWISIDFLAVVPMQSLSWRTLGRASFT--CLAIWALAIAGVVP 236
 QY 243 FVLKQEHVHLVHSEITTCDDVDACSPSPFRFYFVSLAFGLFPLFFVIFCYTTLH 302
 Db 237 LLKQETIQVGLNITTCDDVNETLEGGYAYFSAFSAVF--FFVPLIISTVCYVSIIR 295
 QY 303 KLAS-----KDRIMLGVIKAVL--ILVIFTICEAPTNIILVIHANYHH--NTD 349
 Db 296 CLSSAVANKSKR-----ALFSAVFCIFICFGPTNVLIIAHY-SPLSHSTTE 347
 QY 350 SLVFMVLIALLGLSLNSCLDPFLYFVMS 377
 Db 348 AAFVAILLCVSVSSICCIDPLIYYAS 375

RESULT 14
 PAR4 RAT
 ID PAR4 RAT STANDARD; PRT; 395 AA.
 AC Q920E0;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-
 DE like 3) (Coagulation factor II receptor-like 3).
 GN F2RL3 OR PAR4
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Duoenum;
 RA Chien E.K., Marietti S., Mendoza J., Philippe M.;
 RT "Cloning of the rat protease activated receptor isoforms 3 and 4.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: Receptor for activated thrombin or trypsin coupled to G
 CC proteins that stimulate phosphoinositide hydrolysis. May play a
 CC role in platelets activation.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- PTM: A proteolytic cleavage generates a new amino terminus that
 CC functions as a tethered ligand.
 CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF310216; AAL26790.1; -
 CC EMBL; AF269246; AAK58604.2; -
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm.1; 1
 CC PROSITE; PS00237; G_PROTEIN_RECP_FL1; 1.
 CC PROSITE; PSS0262; G_PROTEIN_RECP_FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Blood coagulation.

FT SIGNAL 1 16
 FT PROPEP 17 58
 FT CHAIN 59 395
 FT DOMAIN 59 93
 FT TRANSMEM 94 114
 FT DOMAIN 115 119
 FT TRANSMEM 120 140
 FT DOMAIN 141 161
 FT TRANSMEM 162 182
 FT DOMAIN 183 203
 FT TRANSMEM 204 224
 FT DOMAIN 225 254
 FT TRANSMEM 255 275
 FT DOMAIN 276 295
 FT TRANSMEM 296 316
 FT DOMAIN 317 330
 FT TRANSMEM 331 354
 FT DOMAIN 355 395
 FT SITE 58 59
 FT DOMAIN 98 101
 FT DISULFID 160 239
 FT CONFLICT 1 9
 FT CONFLICT 192 192
 SQ SEQUENCE 395 AA; 42943 MW; 43B36D0DA22PAPAC CRC64;

Query Match 23.2%; Score 496; DB 1; Length 395;
 Best Local Similarity 31.2%; Pred. No. 3.1e-21;
 Matches 104; Conservative 69; Mismatches 144; Indels 16; Gaps 4;

QY 45 SAKPTLTITKTSFNGGPONTFEEPLSDIEGWTGATTIKAECPEDSISTLHVNNATIGYLR 104
 Db 39 SLRPTVELNESKSPKPNRGP-----GKPCANNSDILELPASSEALL-----LGV-- 85
 QY 105 SSLSTQVPAIYILLFVVGVPNSIVTLMKLSLRKTSISLVIFHNLAIDLPCVTLPK 164
 Db 86 --VPTRLVPAIYGLVWVVGLPANGALWVLAIVRPLPSTILLMNLAVADLLALVLP 143
 QY 165 IAYHLNGNNVFGVMCRITTVFVGNMYCALLITCMGINRYLATAHPFTYQKLPKRSF 224
 Db 144 LVYHLRGQRWPEGEACRVATAALYGHMYGVSLLAAVSLDRYLAIVLSRARALRGQL 203
 QY 225 SLLMCGIWMVFLYMLPFVILKQSVHLVHSEITTCDDVDACSPSPFRFYFVSLAF 284
 Db 204 TTILCLVAVLSATLVLPVLFTRQFLLAGSRMLCHDALPLAEQTSWHR--PAFICLAVL 262
 QY 285 GLIPFVIAIFCYTTLIHKLKSKRIWLGVIKAVLILVIFTICFAPTNIILVIHANY 344
 Db 263 GCFVPLLAMVLCYGATLALAAANGQRYSHAVRLTALVLFSAVAFTPSNVLVLVHYNPS 322
 QY 345 YINTSLSFMYLIALLGLSLNSCLDPFLYFVMS 377
 Db 323 PEAWGNLYGAYVFPVSLALSTLNSCVDPFFIYYVS 355

RESULT 15
 PAR4 MOUSE
 ID PAR4 MOUSE STANDARD; PRT; 396 AA.
 AC O88634;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-
 DE like 3) (Coagulation factor II receptor-like 3).
 GN F2RL3 OR PAR4
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98389762; PubMed=9722561;

FT TRANSMEM 150 168 4 (POTENTIAL).
 FT DOMAIN 169 192 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 193 215 5 (POTENTIAL).
 FT DOMAIN 216 241 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 242 265 6 (POTENTIAL).
 FT DOMAIN 266 287 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 288 312 7 (POTENTIAL).
 FT DOMAIN 313 361 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 313 361 INTERACTION WITH G PROTEINS.
 FT DOMAIN 126 134 BY SIMILARITY.
 FT DISULFID 104 181
 SQ SEQUENCE 361 AA; B5A2171F34C9C67B CRC64;
 Query Match 20.2%; Score 431.5; DB 1; Length 361;
 Best Local Similarity 31.6%; Pred. No. 1.1e-17;
 Matches 95; Conservative 64; Mismatches 107; Indels 35; Gaps 9;
 QY 102 YLRSLSTQVIPAIIYLLFFVGVPSNIVLWKLSTKSI-SLVIFHTNLAIADLLFCVT 160
 DB 24 YAHSTARIWPLHYSLVFTIGLVGNLLALWIVQNRKKNSTLYSTNLVSDILFTTA 83
 QY 161 LPFKIAYHLNNGWVGEVWCRITTVVYGNMCAILLITCMGINRYLATAHPTFYQKLP 220
 DB 84 LPRIAIYAGPDMRGDLCALCRITLVFYNTYAGVNFMTCLSIDRFIAVHPLRYNKK 143
 QY 221 KRFSLLMCGIWMVFLYMLPEVI---LKQEHVHSEITTCDDVDVADCESSPSRPFY 277
 DB 144 RIEHAKGVCFWILVFAQTLLINPMKQE-----AERTCMEYPNEETKS--LPWI 196
 QY 278 FVSLAFGFLIPVPIIFCVTTIHK-----LKSQDIWLGVIKAVLILVIFTIC 328
 DB 197 LLGACFYGVLPIIILICVSOICCKLFTAKONPLTEKSGVNNKALNIIILVIVFLC 256
 QY 329 FAPTNIILVH-----HANY-----YHNTD-SLYEMYLIALCLGSLNCLDPELYFVMS 377
 DB 257 FTFYVAILQHMIKLRFSNLECSQRHSFQISLHF-----TVCLMNFNCMPFIYFPAC 312
 QY 378 K 378
 DB 313 K 313

RESULT 17

P2YR CHICK STANDARD; PRT; 362 AA.
 AC P34956;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
 GN P2Y1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93285340; PubMed=8508924;
 RA Webb T.E., Simon J., Krishek B.J., Bateson A.N., Smart T.G.,
 RA King B.F., Burnstock G., Barnard E.A.;
 RT "Cloning and functional expression of a brain G-protein-coupled ATP
 RT receptor".
 RL FEBS Lett. 324:219-225 (1993).
 RN [2]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=97026278; PubMed=8872457;
 RA van Rhee A.M., Fischer B., van Galen P.J.M., Jacobson K.A.;
 RT "Modeling the P2Y purinoceptor using rhodopsin as template".
 RL Drug Discov. 13:133-140 (1995).
 CC -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
 CC ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN
 CC INSENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT

CC CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, GASTROINTESTINAL TRACT,
 CC SPLEEN AND LEG MUSCLE. IS NOT DETECTED IN THE HEART, LIVER,
 CC STOMACH, LUNG AND KIDNEY.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X73268; CAAS1716.1; .
 CC PIR: S33733; S33733.
 CC PDB: 1DD; 11-JUL-96.
 CC InterPro: IPR00276; GPCR_Rhodpsn.
 CC Pfam: PF00001; 7tm_1; 1.
 CC PRINTS: PR00237; GPCRHOOPS
 CC PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
 CC PROSITE: PS00237; G-PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; 3D-structure.
 KW DOMAIN 1 41
 FT TRANSMEM 42 63 1 (POTENTIAL).
 FT DOMAIN 64 76 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 77 98 2 (POTENTIAL).
 FT DOMAIN 99 115 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 116 136 3 (POTENTIAL).
 FT DOMAIN 137 155 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 156 177 4 (POTENTIAL).
 FT DOMAIN 178 207 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 208 227 5 (POTENTIAL).
 FT DOMAIN 228 254 6 (POTENTIAL).
 FT TRANSMEM 255 274 7 (POTENTIAL).
 FT DOMAIN 275 292 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 293 317 7 (POTENTIAL).
 FT DOMAIN 318 362 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 113 191 BY SIMILARITY.
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 362 AA; 41194 MW; A806C88F89514761 CRC64;
 Query Match 18.7%; Score 399; DB 1; Length 362;
 Best Local Similarity 28.3%; Pred. No. 6.9e-16;
 Matches 91; Conservative 63; Mismatches 122; Indels 46; Gaps 10;
 QY 73 GWTGATTIKACPEDSISTLHVNNATIGYLRSSLSLSTQVIPAIIYLLFFVGVPSNIVLW 132
 DB 21 GWAAGNATTKC-----SLTKGTFQFYVLPVYVILVITGLGNSVAIW 63
 QY 133 KLSLRK---SISLVIFHTNLAIADLLFCVTLPFKIAYHLNNGWVGEVWCRITTVFY 189
 DB 64 MFVFMHWPWSGISVYMF--NLALADFLVLTLPALIFYFNKTDWIFGDMCKLQRFIFH 121
 QY 190 GNMVCAILLTCMGINRYLATAHPTFYQKLPKRSFSLMCGIWMVFLYMLPFIYLKOE 249
 DB 122 VNLGSLILFTICISVHRVTGVVHPLKSLGRLLKKNVYSSVLWALWAVIAP-ILFYSG 180
 QY 250 YHLVHSEITTCDDVDVADCESSPSRPFYFVSLAFGFLIPFVILFCYFTLHLKSKD- 308
 DB 181 TGVRRNKITTCYDDT-ADEYLRSY-FVSMCTTVMFCIPFVILGCGYGLVALLYKDL 238
 QY 309 -----RIMLGVIKAVLILVIFTICFAPNTNIIIVH-HANYYYH-----TDSLIF 353
 DB 239 DNSPLRRKSIYL-----VIIVLTFAVSYPHVMKTLNRLARLDFQTPQMCAFNKVA 293
 QY 354 MYLIACLGSLNSCLDPELYFV 375
 DB 294 TYQVTRGLASLNSCVDPIYFL 315


```
RESULT 18
GP17 HUMAN STANDARD; PRT; 367 AA.
ID Q1304; Q9UDZ6; Q9UE21;
AC 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Probable P2Y purinoceptor GPR17 (P2Y-like receptor) (R12).
GN GPR17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=96145150; PubMed=8558062;
RA Rapoport C.J., Schweickart V.L., Chantry D., Eddy R.L. Jr., Shows T.B.,
RT Godiska R., Gray P.W.;
RT "New members of the chemokine receptor gene family.";
RL J. Leukoc. Biol. 59:18-23(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=hippocampus;
RX MEDLINE=98181695; PubMed=9523551;
RA Blaesus R.H., Weber R.G., Lichter P., Ogilvie A.;
RT "A novel orphan G protein-coupled receptor primarily expressed in the
RT brain is localized on human chromosomal band 2q21.";
RL J. Neurochem. 70:1357-1365(1998).
CC -!- FUNCTION: Putative receptor for purines coupled to G-proteins (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: PRIMARILY EXPRESSED IN BRAIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U33447; AAB16746.1; -
DR EMBL; Y12546; CAA73144.1; -
DR EMBL; Z94154; CAB08107.1; -
DR EMBL; Z94155; CAB08108.1; -
DR HSSP; P34996; 1DDD.
DR Genew; HGNC:4471; GPR17.
DR MIM; 603071; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 85 1 (POTENTIAL).
FT DOMAIN 86 92 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 93 113 2 (POTENTIAL).
FT DOMAIN 114 133 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 134 154 3 (POTENTIAL).
FT DOMAIN 155 175 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 176 196 4 (POTENTIAL).
FT DOMAIN 197 223 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 224 244 5 (POTENTIAL).
FT DOMAIN 245 260 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 261 281 6 (POTENTIAL).
FT DOMAIN 282 308 EXTRACELLULAR (POTENTIAL).
FT
FT TRANSMEM 309 329 7 (POTENTIAL).
FT DOMAIN 330 367 CYTOPLASMIC (POTENTIAL).
FT DISULFID 132 209 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 282 282 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 1 28 MISSING (IN ISOFORM 2).
SQ SEQUENCE 367 AA; 40989 MW; 132FBE97BE83C60C CRC64;
Query Match 18.7%; Score 398.5; DB 1; Length 367;
Best Local Similarity 28.3%; Pred. No. 7.4e-16;
Matches 99; Conservative 60; Mismatches 136; Indels 55; Gaps 9;
QY 42 SNSAKPTLTITKSPNG---GPQWTFEEFPLSDIEGWGTGATTITKACPEDSISTLHVNNA 98
DB 23 SDSS-----QSMNGLEVAPPGLITNFSLATAE-----QCQGE----- 54
QY 99 TIGYLRSLSUSTOVIPIAIYILLFVGVGPSNIVTLWKLRLTKS-LSLIVFHTNLAIADLLF 157
DB 55 -----TPLENMLFASFYLLDFILALVGNLTALWLFIRDHKSGTPTANVFLMHLAVDLSC 108
QY 158 CVTLPEKIAIYHLNGNNWFEVGMCRITTVFVGNMYCAILILTCMGINRYLATAPFTYQ 217
DB 109 VLVLPTRLVYHFGSNHWPGEIACRUTGFLVLMNTASIVYFTICISADRFLAIVHPVKS 168
QY 218 KLPKRSFSLLMCGIVVMVFLYMLPFVILKQYHLVHSEITTTCHDVVDACESPSSPRFY 277
DB 169 KLRRPLIYAHACAFLLVWVAVAPLLVSPQ-----TVQTNHTVVCLQLYREKASHHA 221
QY 278 FVSLAPFGGLIPFVLIIFCYTLIIHLKLS-----KORIWGYKAVLLIIVITFCIPATN 333
DB 222 LVSLA-VAFTFPFITVTTCYLLIIRSLRQGLRVEKRLTKAVRMAIIVLAIFLVCFVPVH 280
QY 334 IILVTHHANYVYHNTDS-----LYFMYLIALGLSINSLCLDPLFLYFVMSK 378
DB 281 VNRSVTLVYHRSGASCATORILANRITSCITSLNGALDPMIFVFAE 330
RESULT 19
ID CLT2_PIG STANDARD; PRT; 345 AA.
AC Q9SN03;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteinyl leukotriene receptor 2 (CysLTR2).
GN CYSLTR2 OR CYSLTR2.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kamohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
RA Ohishi T., Soga T., Matsushime H., Furuichi K.;
RT "Characterization of the cloned rat and porcine cysteinyl leukotriene
RT receptors.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC mediated via a G-protein that activates a phosphatidylinositol-
CC calcium second messenger system (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AB052662; BAB60817.1; -
```



```

Db 122 VNYGSLFCTCSVHRYTGVHPLKSLGRLKKNVYVSSLVWALVAVIAP-ILFYSG 180
QY 250 YHLVHSEITTCVDDVACESPSPFRFYFVSLAFFGLIPFVIIIFCYTTLHKLSKD- 308
Db 181 TGVRRKNTICYTT-ADEVLRYS-FVYSCCTVFNFCIPFVILGCGYGLIVKALYKDL 238
QY 309 -----RIWLGVIKAVLLILVIFTICFAPTNIIILVIH-HANYYHN-----TDSLYF 353
Db 239 DNSPLRKRSIYL-----VIIVLVFAVSYPFPHVMTLNLRLDLFOTPMCAFNDKVA 293
QY 354 MYLIALCLGSLNSCLDPFLYFV 375
Db 294 TYQVTRGLASLNSCVDPILYFL 315

```

RESULT 21

```

P2Y8_XENLA
ID P2Y8_XENLA STANDARD; PRT; 537 AA.
AC P79928;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE P2Y purinoceptor 8 (P2Y8).
GN P2Y8.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Neural plate;
RX MEDLINE=97284734; PubMed=9139711;
RA Bogdanov Y.D., Dale L., King B.F., Whittock N., Burnstock G.;
RT "Early expression of a novel nucleotide receptor in the neural plate
of Xenopus embryos.";
RL J. Biol. Chem. 272:12583-12590(1997).
CC -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ATP, UTP, GTP AND ITP.
CC THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH
CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC MAY PLAY A KEY ROLE IN THE EARLY DEVELOPMENT OF NEURAL TISSUE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99953; CAA68213.1; -.
DR HSPSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
DOMAIN 1 49 EXTRACELLULAR (POTENTIAL).
FT FT 50 70 1 (POTENTIAL).
FT DOMAIN 71 79 CYTOPLASMIC (POTENTIAL).
FT FT 80 100 2 (POTENTIAL).
FT DOMAIN 101 118 EXTRACELLULAR (POTENTIAL).
FT FT 119 139 3 (POTENTIAL).
FT DOMAIN 140 161 CYTOPLASMIC (POTENTIAL).
FT FT 162 182 4 (POTENTIAL).
FT DOMAIN 183 210 EXTRACELLULAR (POTENTIAL).
FT FT 211 231 5 (POTENTIAL).
FT DOMAIN 232 254 CYTOPLASMIC (POTENTIAL).
FT FT 255 275 6 (POTENTIAL).

```

```

FT DOMAIN 276 292 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 293 316 7 (POTENTIAL).
FT DOMAIN 317 537 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 193 BY SIMILARITY.
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 537 AA; 62024 MW; B2CF24812F3C19F2 CRC64;
Query Match 18.4%; Score 393; DB 1; Length 537;
Best Local Similarity 28.5%; Pred. No. 2e-15;
Matches 95; Conservative 62; Mismatches 116; Indels 60; Gaps 11;
QY 87 EDSISTLVNNATIGYLRSSL-----STQ-----VIPAIYILLFVGVSPNSI 128
Db 3 EDIMATSYPTFLTTPYLPKMLNLTNDTDCVDFEGFKLLLPVYSVAVFVGLPLNI 62
QY 129 VTLWKLSLRTKSIIS-LVIEHTNLATADLLFCVTLTPKIAIYHLNGNNNVGEVWMCRTTVV 187
Db 63 AAWWIFIAKORPWNPTTVTFMFLAUSDLYLSLPTLVVYVYADKNWPFGEVLCKLVRFL 122
QY 188 FYGNMYCAIILTCMGINRYLATAHPFTYQKLPKRSFSLLMCGIVWVWVFLVPLFVILK 247
Db 123 FYANLYSSILFTICISVHRVGVCHPITSLRRMNKHAIVICALVWLSVTLCVLPNLI-- 180
QY 248 QBYHLVHSEI--TTCHDVVDACESSSPFRFY--YFVSLAFFGLIPFVIIIFCYTTLIHK 303
Db 181 --FVTVSPKVKNTICHTDT--RPEDFARYVEYSTAIMCLLFGIPCLIIAGCYGLMTRE 234
QY 304 LKS-----KORIWLGVIKAVLLILVIFTICFAPTNIIILVIHANNYY----- 345
Db 235 LMKPIVSGNQOTLPSYKKR-----SIKTIIFVMAFAICMPFPHITRTL---YYARLLG 286
QY 346 ---HNTDLSLYFMYLIALCLGSLNSCLDPFLYFV 375
Db 287 IKCYALNVINVTYKVTTRPLASANSCLDPILYFL 319
RESULT 22
P2Y9_HUMAN
ID P2Y9_HUMAN STANDARD; PRT; 370 AA.
AC Q99677; O15132;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P2Y purinoceptor 9 (P2Y9) (Purinergic receptor 9) (G protein-coupled
receptor GPR23) (P2Y5-like receptor).
GN GPR23 OR P2Y9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97225799; PubMed=9073069;
RA O'Dowd B.F., Nguyen T., Jung B.P., Marchese A., Cheng R.,
RA Heng H.H.Q., Kolakowski L.F. Jr., Lynch K.R., George S.R.;
RT "Cloning and chromosomal mapping of four putative novel human
RT G-protein-coupled receptor genes.";
RL Gene 187:75-81(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97225799; PubMed=9073069;
RA Bohm S.K., Khitin L.M., Payan D.P., Bunnett N.W.;
RP Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97366605; PubMed=9223435;
RA Janssens R., Boeynaems J.M., Godart M., Communi D.;
RT "Cloning of a human heptahelical receptor closely related to the P2Y5
RT receptor.";
RL Biochem. Biophys. Res. Commun. 236:106-112(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: NOT DETECTED IN THE BRAIN REGIONS THALAMUS,
CC PUTAMEN, CAUDATE, FRONTAL CORTEX, PONS, HYPOTHALAMUS, HIPPOCAMPUS.

```

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U65178; AAC51301.1; -
 CC EMBL; U90323; AAB62087.1; -
 CC EMBL; U90322; AAB62088.1; -
 CC EMBL; AF005419; AAB66322.1; -
 CC HSP; P34996; 1DDO.
 CC Genew; HGNC:4478; GPR23.
 CC MIM; 300086; -
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PROSITE; PS00237; G-PROTEIN_RECEPTOR_F1_1; 1.
 CC PROSITE; PS0262; G-PROTEIN_RECEPTOR_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein.
 KW DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 44 64
 FT DOMAIN 65 73
 FT TRANSMEM 74 94
 FT TRANSMEM 95 112
 FT TRANSMEM 113 133
 FT TRANSMEM 134 155
 FT TRANSMEM 156 176
 FT TRANSMEM 177 203
 FT TRANSMEM 204 224
 FT TRANSMEM 225 254
 FT TRANSMEM 255 275
 FT TRANSMEM 276 294
 FT TRANSMEM 295 315
 FT TRANSMEM 316 370
 FT DISULFID 111 188
 FT CARBOHYD 15 15
 FT CARBOHYD 24 24
 FT CARBOHYD 28 28
 FT CARBOHYD 183 183
 FT CONFLICT 192 192
 FT CONFLICT F -> L (IN REF. 3).
 SQ SEQUENCE 370 AA; 41895 MW; 2085752A3929E48 CRC64;
 Query Match 18.1%; Score 387.5; DB 1; Length 370;
 Best Local Similarity 29.3%; Pred. No. 3e-15;
 Matches 93; Conservative 67; Mismatches 112; Indels 45; Gaps 11;
 QY 87 EDSISRL--HVNATIG---YLRSSLSSTQVPIAIYILLFVGVPSNIVTLKLSLRKSI 141
 DB 12 QDSNSSRPRGLGNATANTTCVDSFKYNGAVISVVFILGUTINSVLFVCFRMR 71
 QY 142 S-LVIFHTNLAIADLLFCVTLPEKIAVHLNGNNWFEVGMCRITTVFVGNMYCALLILT 200
 DB 72 SETAIFITNLAVSDLLFVCTLPKIFNFN-RHWPFGDTLCKISGTAFTNIYVGMFLT 130
 QY 201 CMGINRYLATAHPTFYQKLPKRFSLLMCGIWMVFLYMLPVIILKQEVHLVHSITTC 260
 DB 131 CISVDRFLATVFPFRSTIRTRNSAIVCAGVILVLSGGISASLF-----STTNV 181
 QY 261 HDVVDADESSSPRFY--YFVSIAFF---GFLIPFVLIIFCVTLTIHLK----- 305
 DB 182 NNATTCFEGFSKRWKTYLSKTIPIFIEVVGFIPIPLINVCSVVLTRKPKATLSQIG 241
 QY 306 -SKDRILWGIKAVLLIVLTICTFAPTNILVIHH-----ANYYYHNTDSLYPMYLI 357
 DB 242 TNKKV---LKMIVHMAVFCVFPYNSVFLYALVRSQAITNCFLERFAXI--MYPI 295
 QY 358 ALCLGSLNSCLDPFLYF 374
 DB 296 TLCLATLNCDFPIFY 312

RESULT 23
 P2Y5 CHICK STANDARD; PRT; 308 AA.
 ID P2Y5 CHICK
 AC P32250;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE P2Y purinoceptor 5 (P2Y5) (Purinergic receptor 5) (6H1).
 GN P2Y5.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=T-cell;
 RX MEDLINE=93329058; PubMed=8393036;
 RA Kaplan M.H., Smith D.I., Sundick R.S.;
 RT "Identification of a G protein coupled receptor induced in activated
 T cells";
 RL J. Immunol. 151:628-636(1993).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: INDUCED IN ACTIVATED T-CELLS.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L06109; AAB06587.1; -
 CC HSP; P34996; 1DDO.
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PROSITE; PS00237; G-PROTEIN_RECEPTOR_F1_1; 1.
 CC PROSITE; PS0262; G-PROTEIN_RECEPTOR_F1_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
 KW Palmitate.
 FT DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 17 43
 FT DOMAIN 44 52
 FT TRANSMEM 53 76
 FT DOMAIN 77 89
 FT TRANSMEM 90 109
 FT DOMAIN 110 130
 FT TRANSMEM 131 151
 FT DOMAIN 152 178
 FT TRANSMEM 179 206
 FT DOMAIN 207 224
 FT TRANSMEM 225 250
 FT DOMAIN 251 269
 FT TRANSMEM 270 289
 FT DOMAIN 290 308
 FT CARBOHYD 5 5
 FT LIPID 281 281
 FT DISULFID 86 165
 SQ SEQUENCE 308 AA; 35597 MW; 4214E969633B6F7D CRC64;
 Query Match 17.9%; Score 383; DB 1; Length 308;
 Best Local Similarity 30.3%; Pred. No. 4.6e-15;
 Matches 90; Conservative 59; Mismatches 104; Indels 44; Gaps 9;
 QY 115 IYVILFVGVPSNIVTLW---KLSLRKTSISLVIFHTNLAIADLLFCVTLPEKIAHNL 170
 DB 20 VFSMVVFLGLIANCAVIAIFTTLKVRNNTTYML---NLASDLLFVTLFRI-YFV 75
 QY 171 GNNVFGVCMCRITTVFVGNMYCAILLTCMGINRYLATAHPTFYQKLPKRFSLLMCG 230

Db 76 VRNPFQGVLCISVTLFYTNMGSLFLTCISVDRLAIVHPRSKTLRTKKNARIVCV 135

QY 231 IVMVMVFLYMLPFVILKQEHVHLVHSEITTCIDVDVADACESPSSFRFFYFVSLAFF-----GF 286

Db 136 AVMITVLASTPASFTQSTNRQNTQRTCFE-----NPESTWTKYLSRIVFIEIVGF 190

QY 287 LIPFVILIFCYTTLIHLK-----SKRIMWGYIKAVLLILVITTCFAPTNIL 336

Db 191 FIPILINVTCTWVLTANKPLTLRNKLSKKV-----LMIFVHLVIFCFCEVPVYNITL 246

QY 337 VIHHANYYVHTD-----SLYFMYLIALCLGSLNSCLDPFLYFVMSKVVDQLN 384

Db 247 IL-----YSLMRTQTWNCVSVTAIRTMYPVTLCAVSNCCFDEIVYFTSDTNSELD 299

RESULT 24

CLT2 RAT

ID CLT2 RAT STANDARD; PRT; 309 AA.

AC Q924T9;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cysteinyll leukotriene receptor 2 (CysLTR2) (RSBPT32).

GN CysLTR2 OR CysLTR2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA Kamohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,

RA Ohishi T., Soga T., Matsushime H., Furuichi K.;

RT "Characterization of the cloned rat and porcine cysteinyl leukotriene

RL receptors.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Receptor for cysteinyl leukotrienes. The response is

CC mediated via a G-protein that activates a phosphatidylinositol-

CC calcium second messenger system (By similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AB052661; BAB60816.1; -

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7tm_1; 1.

DR PROSITE; PS00237; G-PROTEIN RECP_F1_1; FALSE_NEG.

DR PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein.

FT DOMAIN 1 26 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 27 47 1 (POTENTIAL).

FT DOMAIN 48 56 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 57 77 2 (POTENTIAL).

FT DOMAIN 78 98 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 99 119 3 (POTENTIAL).

FT DOMAIN 120 137 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 138 158 4 (POTENTIAL).

FT DOMAIN 159 187 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 188 208 5 (POTENTIAL).

FT DOMAIN 209 229 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 230 250 6 (POTENTIAL).

FT DOMAIN 251 271 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 272 292 7 (POTENTIAL).

FT DOMAIN 293 309 CYTOPLASMIC (POTENTIAL).

FT DISULFID 95 171 BY SIMILARITY.

FT CARBOHYD 14 14 N-LINKED (GLCNAC...) (POTENTIAL).

FT 166 166 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 167 167 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 309 AA; 35507 MW; A122AC8177879D56 CRC64;

Query Match 17.9%; Score 382.5; DB 1; Length 309;

Best Local Similarity 31.4%; Pred. No. 5e-15;

Matches 93; Conservative 53; Mismatches 113; Indels 37; Gaps 8;

QY 97 NATIGYLRSSLSGTQVPAIYIILLFVGVPSNIVTWLW-LSLRKTSISLVFHTNLAIDL 155

Db 14 NCTI-----ENFKRDFPIYIILFVVGALNGFISVYVFLQYKTSVNVFNLNLAISDF 69

QY 156 LFCVTLPKIAVHLNGNNVWVCEVCRITTVVFGYNNYCAIILITCMGINRVLATAPET 215

Db 70 LFISTLPFRADYNFRGSDWIFGDWACRIMSYSLYNNMTSYIYFLTVLSVRLATAHPFO 129

QY 216 YQKLPRKRSFLMCGIWMVVMVFLYMLPFVILKQEHVHLVHSE-----ITTCIDVDVADACESP 270

Db 130 MLHITSVRSANILCGIIVW-----FIMASSGLLLKHQKKNNTLTCPEL----- 174

QY 271 SSFRFYFVSLAFF---GFLIPFVILIFCYTTLIHLKSKDRIMLG-----YIKAVL 319

Db 175 NLQKFKNLVILNYIALGVGFLLPFFILATCYLLIIRVLLKVEIPESGPRDAQRKALTIV 234

QY 320 LILVIFTTCFAPTNILVIHHANYYHN-TDSLYFMYLIALCLGSLNSCLDPFLYF 374

Db 235 IAMIIIFLCFLPYHALRTIHLVTDADSCMDLHKATVITLTAANSFCNPFLY 290

RESULT 25

CLT2 HUMAN

ID CLT2 HUMAN STANDARD; PRT; 346 AA.

AC Q9NS75; Q9HCQ2;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cysteinyll leukotriene receptor 2 (CysLTR2) (PSEC0146) (HG57) (HPN321).

GN CysLTR2 OR CysLTR2 OR CysLTR2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=20374466; PubMed=10913337;

RA Takasaki J., Kamohara M., Matsumoto M., Saito T., Sugimoto T., Ota T.,

RA Nishikawa T., Kawai Y., Masuho Y., Isogai T., Suzuki Y., Sugano S.,

RA Furuichi K.;

RT "The molecular characterization and tissue distribution of the human

RT cysteinyl leukotriene CysLTR2 receptor.";

RL Biochem. Biophys. Res. Commun. 274:316-322(2000).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=20459128; PubMed=10851239;

RA Heise C.E., O'Dowd B.F., Figueroa D.J., Sawyer N., Nguyen T.,

RA Im D.-S., Stocco R., Bellefeuille J.N., Abramovitz M., Cheng R.,

RA Williams D.L. Jr., Zeng Z., Liu Q., Ma L., Clements M.K., Coulombe N.,

RA Liu Y., Austin C.P., George S.R., O'Neill G.P., Metters K.M.,

RA Lynch K.R., Evans J.F.;

RT "Characterization of the human cysteinyl leukotriene 2 receptor.";

RL J. Biol. Chem. 275:30531-30536(2000).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=20545741; PubMed=11093801;

RA Nothacker H.-P., Wang Z., Zhu Y., Reinscheid R.K., Lin S.H.S.,

RA Civelli O.;

RT "Molecular cloning and characterization of a second human cysteinyl

RT leukotriene receptor: discovery of a subtype selective agonist.";

RN Mol. Pharmacol. 58:1601-1608(2000).

RN [4]

RP SEQUENCE FROM N.A.

RA Dunn M.;

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP MUTAGENESIS OF LYS-107; ARG-110; HIS-262; ARG-265; LYS-289 AND
RP ARG-292
RX MEDLINE=95181393; PubMed=7876172;
RA Erb L., Garrad R.C., Wang Y., Quinn T., Turner J.T., Weisman G.A.;
RT "Site-directed mutagenesis of P2U purinocceptors. Positively charged
amino acids in transmembrane helices 6 and 7 affect agonist potency
and specificity.";
RL J. Biol. Chem. 270:4185-4188 (1995).
CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
activate a phosphatidylinositol-calcium second messenger system.
The affinity range is UTP > ATP > ATP-gamma-S >> 2-methylthio-ATP
= ADP.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: SPLEEN, TESTIS, KIDNEY, LIVER, LUNG, HEART AND
BRAIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L14751; AAA39871.1; -;
DR EMBL; S83099; AAB50735.1; -;
DR EMBL; AK005013; BAB23746.1; -;
DR EMBL; AK017378; BAB30719.1; -;
DR EMBL; BC006613; AAH06613.1; -;
DR HSSP; P34996; 1DDD.
DR MGD; MGI:105107; P2ry2.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR Pfam; PFO0001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOPOPS.
DR PROSITE; PS00237; G PROTEIN RECP Fl_1; 1.
DR PROSITE; PS00262; G PROTEIN RECP Fl_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 33 59 1 (POTENTIAL).
FT DOMAIN 60 70 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 71 93 2 (POTENTIAL).
FT DOMAIN 94 110 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 111 129 3 (POTENTIAL).
FT DOMAIN 130 152 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 153 172 4 (POTENTIAL).
FT DOMAIN 173 194 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 195 220 5 (POTENTIAL).
FT DOMAIN 221 246 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 247 269 6 (POTENTIAL).
FT DOMAIN 270 287 7 (POTENTIAL).
FT TRANSMEM 288 309 7 (POTENTIAL).
FT DOMAIN 310 373 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 9 9 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 106 183 BY SIMILARITY.
FT MUTAGEN 107 107 K->I: NO EFFECT ON RECEPTOR ACTIVATION.
FT MUTAGEN 110 110 R->L: NO EFFECT ON RECEPTOR ACTIVATION.
FT MUTAGEN 262 262 H->L: DECREASE IN RECEPTOR ACTIVATION.
FT MUTAGEN 265 265 R->L: DECREASE IN RECEPTOR ACTIVATION.
FT MUTAGEN 289 289 K->I: NO EFFECT ON RECEPTOR ACTIVATION.
FT MUTAGEN 292 292 R->L: DECREASE IN RECEPTOR ACTIVATION.
FT CONFLICT 17 17 E->D (IN REF. 2).
FT CONFLICT 120 120 S->R (IN REF. 2).
FT CONFLICT 125 125 T->N (IN REF. 2).
FT CONFLICT 196 196 V->M (IN REF. 2).
FT CONFLICT 263 263 V->L (IN REF. 3).
FT CONFLICT 355 355 D->N (IN REF. 2).
FT CONFLICT 369 371 KDI -> PYV (IN REF. 2).

SQ SEQUENCE 373 AA; 42174 MW; 590BBE502E41B3AE CRC64;
Query Match 17.7%; Score 378.5; DB 1; Length 373;
Best Local Similarity 26.9%; Pred. No. 9.6e-15;
Matches 100; Conservative 67; Mismatches 144; Indels 61; Gaps 13;
QY 69 SDIEGWTGATTIKAECPEDSISTLHVNNATGYLRSSLSSTQVIPAIIYILLFVVGVPVSI 128
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 52
3 ADLEPW--NSTINGTWEGDELGYKCRFNEPKYV-----LLPVSYGVVGVGLGCLNV
QY 129 VTLWKLSLRTK---SISLVIPETNLAIADLLFCVTLPPKIAVHLGNMNVFGEVMCRIT 185
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 110
53 VALYIFLCKLKTWNASTYMFH--LAVSDLSLAASLPLLVYYIARGDHPFFSTVLCKLVR
QY 186 VVFYGNMYCAIILATCMGINRYLATAHPFTYOKLPKRSFSLMCGVWVWVFLYMLP--- 242
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 170
111 FLFYTNLYCSILFLACISVHRCLGVLRLPLSLRWGRVARRVAAVWVVLVLAQAPVLY
QY 243 FVLKQEVHLHSEITTHCHVDVVDACESPSSPFYFVSLAFTFGFLIPFVILFIPTTLIH 302
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 223
171 FVTTTS-----VRGTRITCHD-TSARELFSHFVAYSSVLMGLL-FAVPFSVILVYVLMAR
QY 303 KL-----KSKDRIMLGVKAVLLILVIFTICFAPTNIILVIHHA-----NY 344
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 277
224 RLKPAYGTGGLPRAKRS-----VRTIALVAVFALCFLPFHVTRTLTYYSFRSLDLS
QY 345 YHNTDSLYPMYIALCLGSLNSCLDPFLYFVMSKV-----DOLNFXSAMARP----- 392
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 337
278 CHTLNAINWAYKITPLASANSCLDPVLYFLAGQLRVFARDAKPTPTPSQARRKLG
QY 393 LXPRRDIWEDI 404
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 338 LHRPNRTVRKDL 349
RESULT 27
CLT1 HUMAN STANDARD; PRT; 337 AA.
ID CLT1 HUMAN AC Q9Y271;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteinyln leukotriene receptor 1 (CysLTR1) (Cysteinyln leukotriene D4
DE receptor) (LTD4 receptor) (HG55) (HMTMF81).
GN CYSLTR1 OR CYSLTR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RP [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tonsil;
RX MEDLINE=99318129; PubMed=10391245;
RA Lynch K.R., O'Neill G.P., Liu Q., Im D.-S., Sawyer N., Metters K.M.,
RA Coulombe N., Abramovitz M., Figueroa D.J., Zeng Z., Connolly B.M.,
RA Bai C., Austin C.P., Chateaufneuf A., Stocco R., Greig G.M.,
RA Kargman S., Hooks S.B., Hosfield E., Williams D.L. Jr.,
RA Ford-Hutchinson A.W., Caskey C.T., Evans J.F.;
RT "Characterization of the human cysteinyl leukotriene CysLTR1
RT receptor.";
RL Nature 399:789-793 (1999).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood monocytes, Leukocyte, and Spleen;
RX MEDLINE=99393629; PubMed=10462554;
RA Sarau H.M., Ames R.S., Chambers J., Ellis C., Elshourbagy N.,
RA Foley J.J., Schmidt D.B., Muccitelli R.M., Jenkins O., Murdock P.R.,
RA Herrity N.C., Halsey W., Sathe G., Muir A.I., Nuthulaganti P.,
RA Dytko G.M., Buckley P.T., Wilson S., Bergsma D.J., Hay D.W.P.;
RT "Identification, molecular cloning, expression, and characterization
RT of a cysteinyl leukotriene receptor.";
RL Mol. Pharmacol. 56:657-663 (1999).
CC -!- FUNCTION: Receptor for cysteinyl leukotrienes mediating

[illegible]

RESULT 29

	P2YR_RAT	STANDARD;	PRT;	373 AA.
ID	P2YR_RAT			
AC	P49651;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-FEB-1996 (Rel. 33, Last annotation update)			
DE	P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).			
GN	P2RY1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus;			
OC	NCBI TaxID=10116;			

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	EMBL; U22830; AAA91303.1; -	
CC	HSP; P34996; 1DD0	
CC	InterPro; IPR000276; GPCR_Rhodpsn.	
CC	Pfam; PF00001; 7tm_1; 1	
CC	PRINTS; PR00237; GPCRHHODOPSN.	
CC	PROSITE; PS00237; G PROTEIN RECEPT F1 1; 1.	
CC	PROSITE; PS00262; G PROTEIN RECEPT F1 2; 1.	
CC	G-protein coupled receptor; Transmembrane; Glycoprotein.	
CC	DOMAIN 1 52	EXTRACELLULAR (POTENTIAL).
CC	TRANSMEM 1 74	1 (POTENTIAL)
CC	DOMAIN 53 74	CYTOPLASMIC (POTENTIAL).
CC	TRANSMEM 75 87	2 (POTENTIAL).
CC	DOMAIN 88 109	EXTRACELLULAR (POTENTIAL).
CC	TRANSMEM 110 126	3 (POTENTIAL).
CC	DOMAIN 127 147	CYTOPLASMIC (POTENTIAL).
CC	TRANSMEM 148 166	4 (POTENTIAL).
CC	DOMAIN 167 188	EXTRACELLULAR (POTENTIAL).
CC	TRANSMEM 189 218	5 (POTENTIAL).
CC	DOMAIN 219 238	CYTOPLASMIC (POTENTIAL).
CC	TRANSMEM 239 265	6 (POTENTIAL).
CC	DOMAIN 266 285	EXTRACELLULAR (POTENTIAL).
CC	TRANSMEM 286 303	7 (POTENTIAL).
CC	DOMAIN 304 328	CYTOPLASMIC (POTENTIAL).
CC	TRANSMEM 329 373	BY SIMILARITY.
CC	DOMAIN 374 202	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	DISULFID 124 202	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	TRANSMEM 11 11	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	CARBOHYD 27 27	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	TRANSMEM 113 113	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	CARBOHYD 197 197	N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	TRANSMEM 373 AA	6DDF676287B5E648 CRC64;
CC	SEQUENCE 373 AA	423231 MW;

RESULT 30

P2Y ₁ _HUMAN	P2Y ₁ _HUMAN	STANDARD;	PRT;	373 AA.
ID	P2Y1_HUMAN			
AC	P47900;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	13-JUN-2002 (Rel. 41, Last annotation update)			
DE	P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).			
GN	P2RY1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			

RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96257237; PubMed=8666290;
 RA Leon C., Vial C., Cazenave J.-P., Gachet C.;
 RT "Cloning and sequencing of a human cDNA encoding endothelial P2Y1
 RL purinoceptor.";
 RL Gene 171:295-297 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96158962; PubMed=8579591;
 RA Ayyanathan K., Tania W., Harbansjit S., Raghbir A.S., Barnard E.A.,
 RA Kunapuli S.P.;
 RT "Cloning and chromosomal localization of the human P2Y1
 RL purinoceptor.";
 RL Biochem. Biophys. Res. Commun. 218:783-788 (1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96205320; PubMed=8630005;
 RA Janssens R., Communi D., Pirotton S., Samson M., Parmentier M.,
 RA Boeynaems J.M.;
 RT "Cloning and tissue distribution of the human P2Y1 receptor.";
 RL Biochem. Biophys. Res. Commun. 221:588-593 (1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RL Leon C., Vial C., Weber J., Cazenave J.-P., Gacher C.;
 RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 95-373 FROM N.A., FUNCTION, AND INHIBITION STUDIES.
 RC TISSUE=Platelet;
 RX MEDLINE=98113162; PubMed=9442040;
 RA Jin J., Daniel J.L., Kunapuli S.P.;
 RT "Molecular basis for ADP-induced platelet activation. II. The P2Y1
 RL receptor mediates ADP-induced intracellular calcium mobilization and
 RT shape change in platelets.";
 RL J. Biol. Chem. 273:2030-2034 (1998).
 CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
 CC ATP AND ADP. IN PLATELETS BINDING TO ADP LEADS TO MOBILIZATION OF
 CC INTRACELLULAR CALCIUM IONS VIA ACTIVATION OF PHOSPHOLIPASE C. A
 CC CHANGE IN PLATELET SHAPE, AND PROBABLY TO PLATELET AGGREGATION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- INDUCTION: REPPRESSED BY THE P2Y1 RECEPTOR-SPECIFIC ANTAGONISTS
 CC A3PSPS, A3PSP AND A2PSP. THESE INHIBIT CALCIUM ION MOBILIZATION
 CC AND SHAPE CHANGE IN PLATELETS.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; 249205; CAA89066.1; -
 CC EMBL; U42030; AAA97873.1; -
 CC EMBL; U42029; AAA97872.1; -
 CC EMBL; S81950; AAB47091.1; -
 CC EMBL; AJ006945; CAA07339.1; -
 CC EMBL; AF018284; AAB94556.1; -
 CC HSP; P34996; 1DDD.
 CC Genew; HGNC:8539; P2RY1.
 CC MIM; 601167; -
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCRHHODOPSN.
 CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 CC PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Platelet;
 KW Blood coagulation.
 FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 53 74 1 (POTENTIAL).
 FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 88 109 2 (POTENTIAL).
 FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 127 147 3 (POTENTIAL).
 FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 167 188 4 (POTENTIAL).
 FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 219 238 5 (POTENTIAL).
 FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 266 285 6 (POTENTIAL).
 FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 304 328 7 (POTENTIAL).
 FT DOMAIN 329 343 BY SIMILARITY.
 FT DISULFID 124 202 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 138 138 MISSING (IN REF. 1).
 SQ SEQUENCE 373 AA; 42071 MW; 4DC7C668B4145392 CRC64;
 Query Match 17.6%; Score 376.5; DB 1; Length 373;
 Best Local Similarity 27.0%; Pred. No. 1.2e-14; Indels 55; Gaps 12;
 Matches 95; Conservative 72; Mismatches 130;
 QY 46 AKPTLTIKSFNGGQNTPEEFPLSDIEGWTGATTTIKAECPEDSISTLHVNA-TIGYLR 104
 Db 8 AVPNGTDAFLAGPGSS-----WGNSTVA-----STRAVSSSFKALTK 46
 QY 105 SLSLQVTPAIYLLPVGVPSNIVTLWKLSTK---SISLVIFHTNLAIALDLPCVTL 161
 Db 47 TGFQFYYPVAVYILVFIIGLGNVAIWMFVPHMKPWSGISYMF--NLALADFLVLT 104
 QY 162 PKFIAYHLNGNNVGEVWVCRTTIVFYGNMYCAILLTCMGINRYLATAPHTYOKLPK 221
 Db 105 PALIFYYFNKTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVGVYPLKSLGRUK 164
 QY 222 RSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTCDDVVDACESSPSFRFYFVSL 281
 Db 165 KQNAICISVLWLVVVAISP-ILFVSGTGVKKNITCYDT---SDEYLSRFYISM 219
 QY 282 --AFTGFLIPFVILFCYTTLIHKLSKD-----RIWLGYKAVLLILVITFCFA 330
 Db 220 CTTVAMFCVPLVLILGCGYGLVRALYIKDLNSPLRRKSIYL-----VIIIVTFAVSYI 274
 QY 331 PTNIIIVH-HANYVYHN-----TDSLYFMVLIACLGSLNSCLDPLFYFV 375
 Db 275 PFHVMTNLRARLDFOPTAMCAFNDRVYATVQVTRGLASLNSCVDPIILFL 326
 RESULT 31
 P2Y4_MOUSE STANDARD; PRT; 361 AA.
 ID P2Y4_MOUSE AC Q9JUS7;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE P2Y purinoceptor 4 (P2Y4).
 GN P2RY4 OR P2Y4R.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=21185993; PubMed=11290369;
 RA Suarez-Huerta N., Pouillon V., Boeynaems J.-M., Robaye B.;
 RT "Molecular cloning and characterization of the mouse P2Y4 nucleotide
 RL receptor.";
 RL Eur. J. Pharmacol. 416:197-202 (2001).
 CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
 CC activate a phosphatidylinositol-calcium second messenger system.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: Expressed in the liver, intestine, stomach,
 CC bladder and lung.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; AJ277752; CAB91043.1; -;
 DR HSP; P34996; 1DDD.
 DR MGD; MGI:1926594; P2RY4.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 58 1 (POTENTIAL).
 FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 69 91 2 (POTENTIAL).
 FT DOMAIN 92 108 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 109 127 3 (POTENTIAL).
 FT DOMAIN 128 149 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 150 170 4 (POTENTIAL).
 FT DOMAIN 171 192 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 193 218 5 (POTENTIAL).
 FT DOMAIN 219 242 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 243 265 6 (POTENTIAL).
 FT DOMAIN 266 283 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 284 305 7 (POTENTIAL).
 FT DOMAIN 306 361 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 104 181 BY SIMILARITY.
 FT CARBOHYD 175 175 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 361 AA; 41034 MW; 3E8EA84B65BC0A20 CRC64;

 Query Match 17.4%; Score 372.5; DB 1; Length 361;
 Best Local Similarity 27.7%; Pred. No. 2e-14;
 Matches 95; Conservative 61; Mismatches 112; Indels 75; Gaps 14;
 QY 111 VIPAIVILFVGVGVPNSIVTLWKLRLTK---SISLVIFHTNLAIADLLFCVTLPEKIAV 167
 Db 33 LPLSLVAVFVLGLNALNAPTLWLFRLRPWDATYMFH--LALSDTLVLSLPLVYV 90
 QY 168 HLGNNVVFGEVNCRIITVVYVGNMYCAILILTCMGINRYLATAHPTTYOKLPKRSFSL 227
 Db 91 YAAARNHWPFGTGCKFVRFLFYVNLVCSVLFRTCSVHRVYMGICHLPLAIRWGRPRFAGL 150
 QY 228 MCGIVVMVFLYMLPFVILKQEHVHSEIT--CHDVVDACSPSFRFY-YF---VSL 281
 Db 151 LCLGWLWVAGCLVPLNLF---FVTTNANGTTLCHDHT---LPBEFDHYVFSSTIMV 202
 QY 282 AFPGFLIPFVIIIFCYTLLHKL-----KSKDRILWLGVIKAVLLILVIFTCFAPTN 333
 Db 203 LLEFGF--PFLITLVCLVMARLYRPLPGAGQSSR--LRLSLTIAVLTVFVFCVFP-- 256
 QY 334 IILVIHANYYYNHTDSLYFM-----YLIACGLSGLNSCLDPFLY-FV 375
 Db 257 -----FHITRTIYVLABLLNAECRVNLVNVVYKVTPLASANSCLDPVLYLFT 305
 QY 376 MSKVVDOLNXPXAMARLXPRDI-----MEDIH 405
 Db 306 GDKYRNLQOLCRGSTPKRRTTASSLALVTLHBEISRWADIH 348

RESULT 32

P2YR_BOVIN

ID_P2YR_BOVIN STANDARD; PRT; 373 AA.

AC P48042;

DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
 GN P2RY1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aortic endothelium;
 RX MEDLINE=95352058; PubMed=7626079;
 RA Henderson D.J., Elliot D.G., Smith G.M., Webb T.E., Dainty I.A.;
 RT "Cloning and characterisation of a bovine P2Y receptor.";
 RL Biochem. Biophys. Res. Commun. 212:648-656(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Corpus callosum;
 RX MEDLINE=99064562; PubMed=9848096;
 RA Deng G., Matute C., Kumar C.K., Fogarty D.J., Mileti R.;
 RT "Cloning and expression of a P2y purinoceptor from the adult bovine
 corpus callosum.";
 RL Neurobiol. Dis. 5:259-270(1998).
 CC -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
 ATP AND ADP.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X87628; CAA60958.1; -;
 DR EMBL; U34041; AAC78275.1; -;
 DR HSP; P34996; 1DDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 53 74 1 (POTENTIAL).
 FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 88 109 2 (POTENTIAL).
 FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 127 147 3 (POTENTIAL).
 FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 167 188 4 (POTENTIAL).
 FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 219 238 5 (POTENTIAL).
 FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 266 285 6 (POTENTIAL).
 FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 304 328 7 (POTENTIAL).
 FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 124 202 BY SIMILARITY.
 FT CARBOHYD 11 11 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 27 27 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 373 AA; 42287 MW; 9270A7175C0BDA76 CRC64;

 Query Match 17.4%; Score 372; DB 1; Length 373;
 Best Local Similarity 27.8%; Pred. No. 2.2e-14;
 Matches 94; Conservative 63; Mismatches 139; Indels 42; Gaps 10;

QY 59 PONTFEFFPLSDIEGWGATTIKACRPEDSISTLHWNATIGYLRSSLSSTQVIPAIIYL 118
 DB 10 PNCITDTAFLADPSPGNSVTSTAASAPFKALTKTGQFYLL-----PAVYIL 60
 QY 119 LFVVGVPNSIVTLWKLRLTK-----SISLVIFHTNLAIDLFCVTLPPFKIAYHLNGNNVV 175
 DB 61 VFIIGFLGNSVAIWMFVFHMKPSGISVYMF--NLALADFLYVLTLPALIFYFNKTDWI 118
 QY 176 FGEVMCRITTVYGNMYCAILLTTCGNIYRLATAHPTTYQKLPKRSFLLMCGIWMV 235
 DB 119 FGDAMCKLQRFIPHVNLYGSIELTCTISAHRYSGVVYPLKSLGRKKNAVVISVLWLI 178
 QY 236 VFYMLPFLVTKQYEHVHSEIITCDWDVADCSPPSFRFYFVSVL--AFFGFLIPEFVII 293
 DB 179 VVVGISP-ILFYSGTGIRKNTKTCYDTT-----SDEYLSFYFYSCTIVAFECVPLVLI 233
 QY 294 IFCYTTLIHLKSKD-----RIWLGYIKAVLLILVIFTCFAPTNIIILVH-HANY 343
 DB 234 LGCYGLIVRALIYKLDNSPLRKSIVL-----VIVLTVEAVSYIPFHVMTKMLRLARL 288
 QY 344 YVHN-----TDSLFWYLIALCLGSLNSCLDPLFLYFV 375
 DB 289 DFQTPENCAFNDRVYATYQVTRGLASLNSCVDPLIYFL 326
 RESULT 33
 ID_GP34_HUMAN STANDARD; PRT; 381 AA.
 AC Q9UPC5; O95853; PROTEIN RECEPTOR; 381 AA.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable G protein-coupled receptor GPR34.
 GN GPR34.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=99326137; PubMed=10395919;
 RA Schoneberg T., Schulz A., Grosse R., Schade R., Henklein P.,
 RA Schultz G., Gudermann T.;
 RT "A novel subgroup of class I G-protein-coupled receptors.";
 RL Blochim. Biophys. Acta 1446:57-70(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99156852; PubMed=10036181;
 RA Marchese A., Sawzdargo M., Nguyen T., Cheng R., Heng H.H., Nowak T.,
 RA Im D.-S., Lynch K.R., George S.R., O'Dowd B.F.;
 RT "Discovery of three novel orphan G-protein-coupled receptors.";
 RL Genomics 56:12-21(1999).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20434921; PubMed=10982042;
 RA Jacobi F.K., Broghammer M., Pesch K., Zrenner E., Berger W.,
 RA Meindl A., Fusch C.M.;
 RT "Physical mapping and exclusion of GPR34 as the causative gene for
 RT congenital stationary night blindness type 1.";
 RL Hum. Genet. 107:89-91(2000).
 [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahara K., Masuhara Y., Oshima A.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 [5]
 RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ORPHAN RECEPTOR.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: BROADLY EXPRESSED.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AF039686; AAD50531.1; -;
 DR EMBL; AF118670; AAD17248.1; -;
 DR EMBL; AK027780; BAB5362.1; -;
 DR EMBL; BC020678; AAB30678.1; -;
 DR Genew; HGNC:4490; GPR34.
 DR MIM; 300241; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL_1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL_2; 1.
 KW Receptor; G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 61
 FT TRANSMEM 62 88
 FT DOMAIN 83 88
 FT TRANSMEM 89 109
 FT DOMAIN 110 128
 FT TRANSMEM 129 149
 FT DOMAIN 150 171
 FT TRANSMEM 172 192
 FT DOMAIN 193 216
 FT TRANSMEM 217 237
 FT DOMAIN 238 269
 FT TRANSMEM 270 290
 FT DOMAIN 291 310
 FT TRANSMEM 311 331
 FT DOMAIN 332 381
 FT DISULFID 127 204
 FT CARBOHYD 28 28
 FT CARBOHYD 36 36
 FT CARBOHYD 42 42
 FT CARBOHYD 200 200
 FT CARBOHYD 295 295
 FT CONFLICT 181 181
 FT SEQUENCE 381 AA; 43860 MW; 491FC016S5624379 CRC64;
 Query Match 17.2%; Score 367; DB 1; Length 381;
 Best Local Similarity 28.3%; Pred. No. 4.3e-14;
 Matches 105; Conservative 64; Mismatches 130; Indels 72; Gaps 14;
 QY 49 TLTIKSPNGGPONT-----FEFFPLSDIEGWGATTIKACRPEDSISTLHVNATI 100
 DB 7 TMTTTSVSWPYSSHRMFITNHSDDPPQNFSAFNPVT-----CPMD----- 49
 QY 101 GYLRSLSLSTQVIPAIIYLLFVVGVPNSIVTLWKL-LSLRKTSISLVIFHTNLAIDLFCV 159
 DB 50 -----EKLSTVLTTSYSVIFVGLVGNIIALVFLGIHKKRNSIQIYLLNVAIDLILF 105
 QY 160 TLPPKIAVHLNGNNVFGVEVMCRITTVYFVGNMYCAILLTTCGNIYRLATAHPTTYQKL 219
 DB 106 CLPFRIMYHINQNTKLTGVLCKVYVGTFLFYNNMYSIILLGFLISLDRIYKINRSIQORKA 165
 QY 220 PKRSFLLMCGIWMVWV-----FLYMLPFLVTKQYEHVHSEIITCDWDVADCSPPSFRFY 276
 DB 166 ITTKOSIYVCCCIWMALGGFUTMI-ILTLKKGGH-----NSTMCFHYRDKHNAKGAIFN 220
 QY 277 YFVSLAFAFGFLIPFVIIIFCYTTL-----IHLKSK-----DRWLGYIKAVLLILVIFTI 327

Db 221 FILVWFW---LIFLLIILSYIKGNKLLRISRRSKFPNSGKYATTARNSEFVLFIPTI 277
QY 328 CFAPTNIIILVIHANYVYH-----NTDSLYFMYL-----IALCLGSLNSCLDPPFLYFVM 376
Db 278 CFVP-----YHAFPIYISSQLNVSSCYWKEIVHKTNEIMLVLSFNSCLDPPVMYFLM 330
QY 377 S-----KVVDQL 383
Db 331 SSNIRKIMCOL 341

RESULT 34

P2YR_MOUSE
ID P2YR_MOUSE STANDARD; PRT; 373 AA.
AC P49650;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN P2RY1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Insulinoma;
RX MEDLINE=95298025; PubMed=7779087;
RA Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;
RL "Cloning of rat and mouse P2Y purinoceptors";
RN Biochem. Biophys. Res. Commun. 211:211-218(1995).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Leon C.;
RT "Thromboresistance in P2Y1 receptor knockout mice";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC ATP AND ADP. IN PANCREATIC ISLETS, MAY MEDIATE SOME OF THE EFFECTS
CC OF EXTRACELLULAR ATP ON INSULIN SECRETION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; U22829; AAA91302.1; --
DR ENBL; AJ245636; CAB57317.1; --
DR HSDP; P34996; 1DDD.
DR MGSD; MGI:105049; P2RY1.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOODPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPT_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPT_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 53 74 1 (POTENTIAL).
FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 88 109 2 (POTENTIAL).
FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 127 147 3 (POTENTIAL).
FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 167 188 4 (POTENTIAL).
FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 219 238 5 (POTENTIAL).
FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).

FT TRANSSEM 266 285 6 (POTENTIAL).
FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 304 328 7 (POTENTIAL).
FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).
FT DISULFID 124 202 BY SIMILARITY.
FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 373 AA; 42212 MW; 944125E9F4560BB3 CRC64;
Query Match 17.1%; Score 364.5; DB 1; Length 373;
Best Local Similarity 29.5%; Pred. No. 5.7e-14;
Matches 84; Conservative 61; Mismatches 107; Indels 33; Gaps 9;
QY 112 IPAIYILLFVVGVPNSIVTLWKLSLRTK---SISLVIFHTNLAIADLLFCVTLPPKIAIYH 168
Db 54 LPAAVILYFIIGFLGNSVAIMFVFMKPSGISVYMP--NLADFLVILPALIFY 111
QY 169 LGGNNVFEVGMCRITTVVYGNMYCAILILTCMGINRYLATAHPTVYQKLPKRSFSLM 228
Db 112 FNKTDWIFGDAMCKLQRFIFHVNLYGSLFLTCISAHRYSGVVYPLKSLGRLKKNAIVV 171
QY 229 CGIVVMVFLYMLPFVILKQEVHLVHSEITTCDDVDACESSPSRFYFVSL--AFFGF 286
Db 172 SVLVWLVVVAISP-ILFYSGTGTRKNKTVTCYDTT----SNDYLSRYFIYSMCTTVAME 226
QY 287 LIPFVILIFCYTTLTHLKSKD-----RIMLGKAVKALLIIVIFTCFAPTNILV 337
Db 227 CIPLVLILGCGYGLIKVALLYNDLNSPLRRKSIYL-----VIILTVFAVSIPFHVMT 281
QY 338 IH-HANYVYHN-----TDSLYFMYLIALCLGSLNSCLDPPFLYFV 375
Db 282 MNLRARLDPQTPMCDENDRVATYQVTRGLASLNSCVDPILYEL 326

RESULT 35

CLT2_MOUSE
ID CLT2_MOUSE STANDARD; PRT; 309 AA.
AC Q920A1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteinyln leukotriene receptor 2 (CysLTR2).
GN CYSLTR2 OR CYSLTR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=21601669; PubMed=11591709;
RA Hui Y., Yang G., Galczenski H., Figueroa D.J., Austin C.P.,
RA Copeland N.G., Gilbert D.J., Jenkins N.A., Funk C.D.;
RT "The murine cysteinyl leukotriene 2 (CysLTR2) receptor. cDNA and
RT genomic cloning, alternative splicing, and in vitro
RT characterization";
RL J. Biol. Chem. 276:47489-47495(2001).
CC -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC mediated via a G-protein that activates a phosphatidylinositol-
CC calcium second messenger system. The rank order of affinities for
CC the leukotrienes is LTC4 > LTD4 >> LTE4.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed at low levels, with highest
CC expression in the spleen, thymus and adrenal gland, and lower in
CC the kidney, brain and peripheral blood leukocytes.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cysteinylnyl leukotriene receptor 1 (CysLTR1) (Cysteinylnyl leukotriene D4
 DE receptor) (LTD4 receptor)
 GN CysLTR1 OR CysLT1 OR CysLT1R.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC STRAIN=C57BL/6; TISSUE=Trachea;
 RX PubMed=11226226;
 RA Maekawa A., Kanaoka Y., Lam B.K., Austen K.F.;
 RT "Identification in mice of two isoforms of the cysteinyl leukotriene 1
 RT receptor that result from alternative splicing.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:2256-2261(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=C57BL/6; TISSUE=Trachea;
 RX MEDLINE=21562332; PubMed=11705452;
 RA Martin V., Sawyer N., Stocco R., Unett D., Lerner M.R., Abramovitz M.,
 RA Funk C.D.;
 RT "Molecular cloning and functional characterization of murine
 RT cysteinyl-leukotriene 1 (CysLT1) receptors.";
 RL Biochem. Pharmacol. 62:1193-1200(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Blood;
 RX MEDLINE=21240338; PubMed=11342226;
 RA Mollerup J., Jorgensen S.T., Hougard C., Hoffmann E.K.;
 RT "Identification of a murine cysteinyl leukotriene receptor by
 RT expression in Xenopus laevis oocytes.";
 RL Biochim. Biophys. Acta 1517:455-459(2001).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=T-cell;
 RX Ogasawara H., Izumi T., Shimizu T.;
 RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: Receptor for cysteinyl leukotrienes mediating
 CC constriction of the microvascular smooth muscle during an
 CC inflammatory response. This response is mediated via a G-protein
 CC that activates a phosphatidylinositol-calcium second messenger
 CC system. The rank order of affinities for the leukotrienes is LTD4
 CC >> LTE4 = LTC4 >> LTB4.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- ALTERNATIVE PRODUCTS: 2 isoforms; 1/long form (shown here) and
 CC 2/short form; are produced by alternative splicing.
 CC -I- TISSUE SPECIFICITY: Widely expressed, with higher expression in
 CC the lung and skin, intermediate levels in the heart, kidney and
 CC stomach and lower levels in several other tissues. Isoform 1 is
 CC the most abundant form in all tested tissues.
 CC -I- MISCELLANEOUS: MK-571, a selective antagonist, was shown to
 CC inhibit eosinophilia, bronchial hyperreactivity and microvascular
 CC leakage. Zafirlukast (Accolate) and pranlukast (Onon) were also
 CC shown to be selective antagonists.
 CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF329272; AA16715.1; -
 CC EMBL; AF329272; AA16716.1; -
 CC EMBL; AF205830; AA15433.1; -
 CC EMBL; AF263370; AA73047.1; -
 CC EMBL; AB044087; BA96809.1; -
 CC HSP; P34996; 1DDD.
 DR MGD; MGI:1926218; Cysltr1.

DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN RECP_F1_1; FALSE_NEG.
 DR PROSITE; PS00262; G-PROTEIN RECP_F1_2; 1. Glycoprotein;
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Alternative splicing.
 FT DOMAIN 1 43
 FT TRANSMEM 44 64
 FT DOMAIN 65 72
 FT TRANSMEM 73 93
 FT DOMAIN 94 121
 FT TRANSMEM 122 142
 FT DOMAIN 143 156
 FT TRANSMEM 157 177
 FT DOMAIN 178 208
 FT TRANSMEM 209 229
 FT DOMAIN 230 245
 FT TRANSMEM 246 266
 FT DOMAIN 267 291
 FT TRANSMEM 292 312
 FT DOMAIN 313 352
 FT DISULFID 111 188
 FT CARBOHYD 15 15
 FT CARBOHYD 19 19
 FT CARBOHYD 26 26
 FT CARBOHYD 184 184
 FT VARSPLIC 1 13
 FT CONFLICT 176 176
 FT SEQUENCE 352 AA; 40715 MW; 5BDC94B3F1CD0CAB CRC64;
 SQ
 Query Match 16.9%; Score 361; DB 1; Length 352;
 Best Local Similarity 30.8%; Pred. No. 8.6e-14;
 Matches 103; Conservative 58; Mismatches 129; Indels 44; Gaps 13;
 QY 58' GPQNTFEEFPLSDIEGWTGATTIKAECPDSISTLVHVNNA---TIGYLRSSLSTQVIPA 114
 DB 5 GTKQTF---LENNMGTEINLTSL-----INNTCHDTIDEFRN---QVYST 43
 QY 115 IYVLLFVVGVPSPNIVTLWKLSLRT--KSLVLIPTHTNLAIALDLFCVTLPPKIAVHLGN 172
 DB 44 MYSVLSVVGFFGNSVLYVL-IKTHKSAFOVMINLAIALDLCCVTLPLRVVYVHKG 102
 QY 173 NMVFGVMCRITTVVYFGNMYCAILILTCMGINRYLATAHPPTQKLPKRSFLMCGIV 232
 DB 103 KWLFGDFLCRLTYALYNLYCSIFFTAMSFRCVAIVFPQINLVTKKARFVIGI 162
 QY 233 WYVFLYMLPFVILKQVHVLHSEITTCDDVVVDACESPSFRFYFVSLAFPGFLIPFVI 292
 DB 163 WIFVILTSSPFLMYKS--YQDEKNNTKCFEPQNNQAKYVLLHVVSL-FFGFIPFVT 219
 QY 293 IIFCYTTLIHLK--KSKDRILWLYIKAVLLILVI---FTICFAPTNILVTHHANYHN 347
 DB 220 IIVCYVTMIITLLKNTKKNMPSRKATGMITVTAALFVSPFYHIQRTIH--LHLHS 277
 QY 348 ----TDSLVFM---YLIALCLGSLNSCLDPFLYF 374
 DB 278 ETRPCDVLRLMOKSVVITLSLAASCCFDPFLYF 311
 RESULT 39
 GP34_MOUSE
 ID GP34_MOUSE STANDARD; PRT; 375 AA.
 AC Q9RIK6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable G protein-coupled receptor GPR34.
 GN GPR34.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:01:21 ; Search time 17.9779 Seconds
(without alignments)
666.102 Million cell updates/sec

Title: US-09-208-629F-3

Perfect score: 2136
Sequence: 1 TLVTXQHPVAGSQDIKWIL.....AMARPLXRRPRDIWIDHAW 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2130	99.7	407	2	US-08-742-440A-3
2	1433.5	67.6	408	2	US-08-742-440A-6
3	581.5	27.2	395	1	US-08-097-938-5
4	581.5	27.2	395	1	US-08-476-000-5
5	581.5	27.2	395	1	US-08-472-840-5
6	581.5	27.2	395	2	US-08-476-976-5
7	581.5	27.2	395	3	US-08-474-410-5
8	581.5	27.2	395	4	US-08-486-673B-2
9	581.5	27.2	395	4	US-08-486-673B-5
10	581.5	27.2	399	1	US-08-476-000-61
11	581.5	27.2	399	1	US-08-472-840-61
12	581.5	27.2	399	2	US-08-476-976-61
13	581.5	27.2	399	3	US-08-474-410-61
14	581.5	27.2	399	4	US-08-486-673B-61
15	579.5	27.1	395	1	US-08-097-938-2
16	579.5	27.1	395	1	US-08-476-000-2
17	579.5	27.1	395	1	US-08-472-840-2
18	579.5	27.1	395	2	US-08-476-976-2
19	579.5	27.1	395	3	US-08-474-410-2
20	563	26.4	398	1	US-08-097-938-4
21	563	26.4	398	1	US-08-476-000-4
22	563	26.4	398	1	US-08-472-840-4
23	563	26.4	398	2	US-08-476-976-4
24	563	26.4	398	3	US-08-474-410-4
25	563	26.4	398	4	US-08-486-673B-4
26	561.5	26.3	398	4	US-08-486-673B-6
27	560.5	26.2	398	1	US-08-097-938-6

28	560.5	26.2	398	1	US-08-476-000-6	Sequence 6, Appli
29	560.5	26.2	398	1	US-08-472-840-6	Sequence 6, Appli
30	560.5	26.2	398	2	US-08-476-976-6	Sequence 6, Appli
31	560.5	26.2	398	3	US-08-474-410-6	Sequence 6, Appli
32	559	26.2	397	4	US-08-486-673B-63	Sequence 63, Appli
33	553.5	25.9	394	2	US-08-742-440A-8	Sequence 8, Appli
34	551	25.8	397	1	US-08-476-000-63	Sequence 63, Appli
35	551	25.8	397	1	US-08-472-840-63	Sequence 63, Appli
36	551	25.8	397	2	US-08-476-976-63	Sequence 63, Appli
37	551	25.8	397	3	US-08-474-410-63	Sequence 63, Appli
38	509	23.8	425	1	US-07-657-769B-69	Sequence 69, Appli
39	509	23.8	425	1	US-07-789-184-220	Sequence 220, App
40	509	23.8	425	1	US-08-475-263-220	Sequence 220, App
41	509	23.8	425	1	US-08-485-886-220	Sequence 220, App
42	509	23.8	425	2	US-08-477-362-220	Sequence 220, App
43	509	23.8	425	2	US-08-477-134-220	Sequence 220, App
44	509	23.8	425	2	US-08-911-320A-3	Sequence 3, Appli
45	509	23.8	425	2	US-08-742-440A-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-742-440A-3
; Sequence 3, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 407 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; US-08-742-440A-3
Query Match 99.7%; Score 2130; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 3.3e-155;

Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLYXQHPVAGSODIKMKILILVAAGLLFLPVTVCOSGINVSDNSAKPTLTIKSFNGGPQ 60
Db 1 TLYXQHPVAGSODIKMKILILVAAGLLFLPVTVCOSGINVSDNSAKPTLTIKSFNGGPQ 60

Qy 61 NTFFEFPLSDIEGWTGATTTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLF 120
Db 61 NTFFEFPLSDIEGWTGATTTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLF 120

Qy 121 VVGVPNSIVTLWKLRLTKSLVIFHTNLAIADLLFCVTLPFKIAHLNNGNWWFGEVM 180
Db 121 VVGVPNSIVTLWKLRLTKSLVIFHTNLAIADLLFCVTLPFKIAHLNNGNWWFGEVM 180

Qy 181 CRITTVVFGNMYCAIILITCMGINRYLATAHPTYOQLPKRSFSLMCGIIVWVWFLYM 240
Db 181 CRITTVVFGNMYCAIILITCMGINRYLATAHPTYOQLPKRSFSLMCGIIVWVWFLYM 240

Qy 241 LPFVLKQEHVHSEITTCDDVDACESPSPFRFYFVSLAFPGFLIPFVIIIFCVTTL 300
Db 241 LPFVLKQEHVHSEITTCDDVDACESPSPFRFYFVSLAFPGFLIPFVIIIFCVTTL 300

Qy 301 IHKLKSDRIWLGVIKAVLLILVIFTICFAPTNIILVIHHANYHHNTDSLYFMYLIAC 360
Db 301 IHKLKSDRIWLGVIKAVLLILVIFTICFAPTNIILVIHHANYHHNTDSLYFMYLIAC 360

Qy 361 LGSLSCLDPLFYFMSKVVDQLNPXSAMARPLXPRRDIWEDIHAW 407
Db 361 LGSLSCLDPLFYFMSKVVDQLNPXSAMARPLXPRRDIWEDIHAW 407

RESULT 2

US-08-742-440A-6

; Sequence 6, Application US/08742440A

; Patent No. 5892014

; GENERAL INFORMATION:

; APPLICANT: Coughlin, Shaun

; APPLICANT: Ishihari, Hiroaki

; APPLICANT: Connolly, Andrew

; TITLE OF INVENTION: Protease Activated Receptor

; TITLE OF INVENTION: 3 and Uses Thereof

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bozicevic & Reed, LLP

; STREET: 285 Hamilton Avenue, Suite 200

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/742,440A

; FILING DATE: 30-OCT-1996

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Sherwood, Pamela J

; REGISTRATION NUMBER: 36,677

; REFERENCE/DOCKET NUMBER: UCAL/060PAT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-327-3400

; TELEFAX: 650 327-3231

; TELEX:

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 408 amino acids

; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
US-08-742-440A-6

Query Match 67.6%; Score 1443.5; DB 2; Length 408;

Best Local Similarity 71.5%; Pred. No. 7.5e-103;

Matches 264; Conservative 45; Mismatches 59; Indels 1; Gaps 1;

Qy 11 GSODIKMKILILVAAGLLFLPVTVCOSGI-NVSDNSAKPTLTIKSFNGSPONTFFEEFPLS 69
Db 14 GTQVIRKALIFAAGLLLLPTFCOSGMENDTNLAKPTLPKIKTFRGAAPPNSFEFPFS 73

Qy 70 DIEGTGATTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIV 129
Db 74 ALEGWTGATTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPANAV 133

Qy 130 TLMKLSLRKTSISLVIFHTNLAIADLLFCVTLPFKIAHLNNGNWWFGEVMCRITTVFY 189
Db 134 TLMWLFRTSICCTTVFYTNLAIDFLCVTLPFKIAHLNNGNWWFGEVLCRATTIVFY 193

Qy 190 GNNYCAIILITCMGINRYLATAHPTYOQLPKRSFSLMCGIIVWVWFLYMLPFVILKQE 249
Db 194 GNNYCSILLIACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFVILKQE 253

Qy 250 YHLVHSEITTCDDVDACESPSPFRFYFVSLAFPGFLIPFVIIIFCYTTLIHKLKSKOR 309
Db 254 YVLPQDITTCDDVHNTCESSPFQYFISLAFPGFLIPFVIIIFCYTTLIHKLKSKOR 313

Qy 310 IWLGYIKAVLLILVIFTICFAPTNIILVIHHANYHHNTDSLYFMYLIACLSGLNSCLD 369
Db 314 RMLWYVKASILLIVIFTICFAPSNIILVIHHANYHHNTDGLYFIYLIACLSGLNSCLD 373

Qy 370 PFLYFWSK 378
Db 374 PFLYFLMSK 382

RESULT 3

US-08-097-938-5

; Sequence 5, Application US/08097938

; Patent No. 5629174

; GENERAL INFORMATION:

; APPLICANT: SUNDELIN, JOHAN

; APPLICANT: SCARBOROUGH, ROBERT M.

; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS

; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS

; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

; CITY: Washington, D.C.

; COUNTRY: USA

; ZIP: 20006-1812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/097,938

; FILING DATE: 26-JUL-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: MURASHIGE, KATE H.

; REGISTRATION NUMBER: 29,959

; REFERENCE/DOCKET NUMBER: 22803-20006.00

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 887-1500

; TELEFAX: (202) 887-0763

; TELEX: 90-4030

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 395 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-097-938-5

Query Match      27.2%; Score 581.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY 95 VNNATIGYLRSSLSQVIPAIIYLLFVVGVPSTVTLWKLRLT-KSISLVIFHTNLAI 153
Db 59 IDEFSASILTKLTTFVLPVYIIIVFVIGLPSNGMALWIFLRTKKKHPAVIYMANLALA 118
QY 154 DLLFCVTLPEKTAIYHLNGNNWVGEVWCRITTVVFGNNMYCAIILITCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
QY 214 FTYQKLPRKSFSLMCGIWMVFLYMLPFLVILKQEVHLVHSEITTTCHDVVDACESPSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYPALNITTTCHDVLP-EVLVGD 236
QY 274 RYFYVSLAFLGFLIPFVILFCYTTLI-----HKLSKDRIMWLGVIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA-----IRLIITVLA 292
QY 324 IFTICFAPTNILVIHHANYHYHNTDSLYFMYLIALCLGSLNSCLDPLFYFVMSK 378
Db 293 MYFICFAPSNNLLLVHVFILIKTORQSHVYALYLVALCLSTLNSCIDPFFVYFVSK 347

RESULT 4
US-08-476-000-5
; Sequence 5, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-476-000-5
```

```

; TOPOLOGY: linear
US-08-476-000-5

Query Match      27.2%; Score 581.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

QY 95 VNNATIGYLRSSLSQVIPAIIYLLFVVGVPSTVTLWKLRLT-KSISLVIFHTNLAI 153
Db 59 IDEFSASILTKLTTFVLPVYIIIVFVIGLPSNGMALWIFLRTKKKHPAVIYMANLALA 118
QY 154 DLLFCVTLPEKTAIYHLNGNNWVGEVWCRITTVVFGNNMYCAIILITCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
QY 214 FTYQKLPRKSFSLMCGIWMVFLYMLPFLVILKQEVHLVHSEITTTCHDVVDACESPSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYPALNITTTCHDVLP-EVLVGD 236
QY 274 RYFYVSLAFLGFLIPFVILFCYTTLI-----HKLSKDRIMWLGVIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA-----IRLIITVLA 292
QY 324 IFTICFAPTNILVIHHANYHYHNTDSLYFMYLIALCLGSLNSCLDPLFYFVMSK 378
Db 293 MYFICFAPSNNLLLVHVFILIKTORQSHVYALYLVALCLSTLNSCIDPFFVYFVSK 347

RESULT 5
US-08-472-840-5
; Sequence 5, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-472-840-5
```

[illegible]

RESULT 6
 US-08-476-976-5
 Sequence 5, Application US/08476976
 Patent No. 5874400
 GENERAL INFORMATION:
 APPLICANT: SUNDELIN, JOHAN
 APPLICANT: SCARBOROUGH, ROBERT M.
 TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 NUMBER OF SEQUENCES: 63
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476,976
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/390,301
 FILING DATE: 25-JAN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: ADLER, REID G.
 REGISTRATION NUMBER: 30,988
 REFERENCE/DOCKET NUMBER: 2803-0006.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 TELEX: 90-4030
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 395 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-476-976-5

Query Match	27.2%	Score 581.5;	DB 2;	Length 395;
Best Local Similarity	37.6%	Pred. No. 4e-37;		
Matches 111; Conservative	65;	Mismatches 102;	Indels 17;	Gaps 5;

[illegible]

RESULT 7
US-08-474-410-5
; Sequence 5, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:

CLASSIFICATION: ;
 PRIOR APPLICATION DATA: US 08/390,301 ;
 APPLICATION NUMBER: US 08/390,301 ;
 FILING DATE: 25-JAN-1995 ;
 ATTORNEY/AGENT INFORMATION: ;
 NAME: ADLER, REID G. ;
 REGISTRATION NUMBER: 30,988 ;
 REFERENCE/DOCKET NUMBER: 2803-0006.20 ;
 TELECOMMUNICATION INFORMATION: ;
 TELEPHONE: (202) 887-1500 ;
 TELEFAX: (202) 887-0763 ;
 TELEX: 90-4030 ;
 INFORMATION FOR SEQ ID NO: 5: ;
 SEQUENCE CHARACTERISTICS: ;
 LENGTH: 395 amino acids ;
 TYPE: amino acid ;
 STRANDEDNESS: single ;
 TOPOLOGY: linear ;
 US-08-474-410-5 ;

Query Match 27.2%; Score 581.5; DB 3; Length 395;
Best Local Similarity 37.6%; Pred. No. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

Qy .95 VNNATTGYLRSSLSGTVPAYIYLLFVGVGPSNIVTLWKLSLRT-KSISLVIFHTNLAIA 153
:: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: :

Db 59 IDEFSASILTKLTTFVLPVYIIIVFVIGLPSNGMALWIFLFRKKKHPAVIYMANLALA 118
Qy 154 DLLFCVTLPPKIAHYHLNGNNWVGEVNCRIITVVYFGNMYCAILILTCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLHNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
Qy 214 FTYQKLPRFSLSLMCGIWMVWFLYMLPFVILKQEHVHSEITTCDDVVDACESSPSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTCDDVLP-EVLVGD 236
Qy 274 RFYFVSLAFPGFLIPFVIIIFCYTTLI-----HKLKSKDRILWGIYKAVLLILV 323
Db 237 MFNYFLSLATGVLPFALLTASAYVLMIKTLRSSAMDEHSEKKRQRA----IRLIITVLA 292
Qy 324 IFTICFAPTNIIIVHANYHHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 293 MYFICFAPSNNLLVVHVFILIKTORQSHVYALYLVALCLSTLNSCIDPFVYFVYFSK 347

RESULT 8
US-08-486-673B-2
; Sequence 2, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 395
; ORGANISM: Mus musculus
US-08-486-673B-2

Query Match 27.2%; Score 581.5; DB 4; Length 395;
Best Local Similarity 37.6%; Pred. No. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

Qy 95 VNNATICYLRSSLSQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTKLTTFVLPVYIIIVFVIGLPSNGMALWIFLFRKKKHPAVIYMANLALA 118
Qy 154 DLLFCVTLPPKIAHYHLNGNNWVGEVNCRIITVVYFGNMYCAILILTCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLHNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
Qy 214 FTYQKLPRFSLSLMCGIWMVWFLYMLPFVILKQEHVHSEITTCDDVVDACESSPSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTCDDVLP-EVLVGD 236
Qy 274 RFYFVSLAFPGFLIPFVIIIFCYTTLI-----HKLKSKDRILWGIYKAVLLILV 323
Db 237 MFNYFLSLATGVLPFALLTASAYVLMIKTLRSSAMDEHSEKKRQRA----IRLIITVLA 292
Qy 324 IFTICFAPTNIIIVHANYHHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 293 MYFICFAPSNNLLVVHVFILIKTORQSHVYALYLVALCLSTLNSCIDPFVYFVYFSK 347

RESULT 9
US-08-486-673B-5
; Sequence 5, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan

; APPLICANT: Scarborough, Robert M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: Protein sequence of C140 receptor
US-08-486-673B-5

Query Match 27.2%; Score 581.5; DB 4; Length 395;
Best Local Similarity 37.6%; Pred. No. 4e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;
Qy 95 VNNATICYLRSSLSQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTKLTTFVLPVYIIIVFVIGLPSNGMALWIFLFRKKKHPAVIYMANLALA 118
Qy 154 DLLFCVTLPPKIAHYHLNGNNWVGEVNCRIITVVYFGNMYCAILILTCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLHNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
Qy 214 FTYQKLPRFSLSLMCGIWMVWFLYMLPFVILKQEHVHSEITTCDDVVDACESSPSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMKQTIYIPALNITTCDDVLP-EVLVGD 236
Qy 274 RFYFVSLAFPGFLIPFVIIIFCYTTLI-----HKLKSKDRILWGIYKAVLLILV 323
Db 237 MFNYFLSLATGVLPFALLTASAYVLMIKTLRSSAMDEHSEKKRQRA----IRLIITVLA 292
Qy 324 IFTICFAPTNIIIVHANYHHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 293 MYFICFAPSNNLLVVHVFILIKTORQSHVYALYLVALCLSTLNSCIDPFVYFVYFSK 347

RESULT 10
US-08-476-000-61
; Sequence 61, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995

ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-000-61

Query Match 27.2%; Score 581.5; DB 1; Length 399;
Best Local Similarity 37.6%; Pred. No. 4.1e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;
QY 95 VNNATIGVRLSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 63 IDEFSASILTGLTIVFLPVVYIIIVFVIGLPSNGMALWIFLFRKKKHPAVIYMANLALA 122
QY 154 DLLFCVTLPEFKIAYHLGNWVGFGEVWCRITTVVFGNMYCAIILITCMGINRYLATAHP 213
Db 123 DLLSVIWFPLKISYHLGNWVGEALCKVLIGFYGNNYCSILFMTCLSVQRVYVWVNP 182
QY 214 FTYOKLPKRSFSLMCGIWMVFLYMLPFVILKOEYHLVHSEITTCCHDVVDACESPSSF 273
Db 183 MGHPR-KKANIAGVSLAIWLLIFLVTIPLVYMKQTIYIPALNITTCCHDVLP-EVLVGD 240
QY 274 RFYFVSLAPFGFLIPFVIIIFCYTTLI-----HKLKSKDRIWGLYKAVILLIV 323
Db 241 MENYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA---IRLIITVLA 296
QY 324 IFTICFAPTNILVIHANYHYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMSK 378
Db 297 MYFICFAPSNNLLVHVHFLIKTORQSHVYALYVALCLSTLNSCIDPFFVYFVSK 351

RESULT 11
US-08-472-840-61
Sequence 61, Application US/08472840
Patent No. 5763575
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472.840
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/390.301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: 2803-0006.20
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-840-61

Query Match 27.2%; Score 581.5; DB 1; Length 399;
Best Local Similarity 37.6%; Pred. No. 4.1e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;
QY 95 VNNATIGVRLSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 63 IDEFSASILTGLTIVFLPVVYIIIVFVIGLPSNGMALWIFLFRKKKHPAVIYMANLALA 122
QY 154 DLLFCVTLPEFKIAYHLGNWVGFGEVWCRITTVVFGNMYCAIILITCMGINRYLATAHP 213
Db 123 DLLSVIWFPLKISYHLGNWVGEALCKVLIGFYGNNYCSILFMTCLSVQRVYVWVNP 182
QY 214 FTYOKLPKRSFSLMCGIWMVFLYMLPFVILKOEYHLVHSEITTCCHDVVDACESPSSF 273
Db 183 MGHPR-KKANIAGVSLAIWLLIFLVTIPLVYMKQTIYIPALNITTCCHDVLP-EVLVGD 240
QY 274 RFYFVSLAPFGFLIPFVIIIFCYTTLI-----HKLKSKDRIWGLYKAVILLIV 323
Db 241 MENYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA---IRLIITVLA 296
QY 324 IFTICFAPTNILVIHANYHYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMSK 378
Db 297 MYFICFAPSNNLLVHVHFLIKTORQSHVYALYVALCLSTLNSCIDPFFVYFVSK 351

RESULT 12
US-08-476-976-61
Sequence 61, Application US/08476976
Patent No. 5874400
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476.976
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELEPHONE: (202) 887-1500


```
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-61

Query Match      27.2%; Score 581.5; DB 2; Length 399;
Best Local Similarity 37.6%; Pred. No. 4.1e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 63 IDEFSASILTGLTITVFLPVYIIVFVIGLPSNGMALWIFLRTKKKHPAVIYMANLALA 122

Qy 154 DLLFCVTLTPFKIAYHLNGNNWVGEVWCRITTVVFGYNNMYCAIILITCMGINRYLATAHP 213
Db 123 DLLSVIWFPLKISYHLHNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182

Qy 214 FTYOKLPKRSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTCDDVVDACESPSF 273
Db 183 MGHPR-KKANIAGVSLAIWLILFLVTIPLYVMKQTIYPALNITTCDDVLP-EVLVGD 240

Qy 274 RFYFVSLAFPGFLIPFVIIICYTTLI-----HKLKSKDRIWLGVIKAVLLILV 323
Db 241 MFNYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA-----IRLIITVLA 296

Qy 324 IFTICFAPNTIILVHHANYHHNTDSLYFMYLIALCLGSLNSCLDPFLFYVMSK 378
Db 297 MYFICFAPSLLLVVHYFLIKTORQSHVYALYVALCLSTLNSCIDPFPVYFVSK 351

RESULT 13
US-08-474-410-61
; Sequence 61, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 61:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-410-61

Query Match      27.2%; Score 581.5; DB 3; Length 399;
Best Local Similarity 37.6%; Pred. No. 4.1e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 63 IDEFSASILTGLTITVFLPVYIIVFVIGLPSNGMALWIFLRTKKKHPAVIYMANLALA 122

Qy 154 DLLFCVTLTPFKIAYHLNGNNWVGEVWCRITTVVFGYNNMYCAIILITCMGINRYLATAHP 213
Db 123 DLLSVIWFPLKISYHLHNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182

Qy 214 FTYOKLPKRSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTCDDVVDACESPSF 273
Db 183 MGHPR-KKANIAGVSLAIWLILFLVTIPLYVMKQTIYPALNITTCDDVLP-EVLVGD 240

Qy 274 RFYFVSLAFPGFLIPFVIIICYTTLI-----HKLKSKDRIWLGVIKAVLLILV 323
Db 241 MFNYFLSLAIGVFLPALLTASAYVLMIKTLRSSAMDEHSEKKRQRA-----IRLIITVLA 296

Qy 324 IFTICFAPNTIILVHHANYHHNTDSLYFMYLIALCLGSLNSCLDPFLFYVMSK 378
Db 297 MYFICFAPSLLLVVHYFLIKTORQSHVYALYVALCLSTLNSCIDPFPVYFVSK 351

RESULT 14
US-08-486-673B-61
; Sequence 61, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Mus musculus
; US-08-486-673B-61

Query Match      27.2%; Score 581.5; DB 4; Length 399;
Best Local Similarity 37.6%; Pred. No. 4.1e-37;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSSTQVIPAIYILLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 63 IDEFSASILTGLTITVFLPVYIIVFVIGLPSNGMALWIFLRTKKKHPAVIYMANLALA 122

Qy 154 DLLFCVTLTPFKIAYHLNGNNWVGEVWCRITTVVFGYNNMYCAIILITCMGINRYLATAHP 213
Db 123 DLLSVIWFPLKISYHLHNGNNWVGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182

Qy 214 FTYOKLPKRSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTCDDVVDACESPSF 273
Db 183 MGHPR-KKANIAGVSLAIWLILFLVTIPLYVMKQTIYPALNITTCDDVLP-EVLVGD 240

Qy 274 RFYFVSLAFPGFLIPFVIIICYTTLI-----HKLKSKDRIWLGVIKAVLLILV 323
```

Db 241 MFNYFLSLAIGVFLFPALLTASAYVLMKILTRSSAMDEHSEKQRRA-----IRLIITVLA 296
Qy 324 IFTICFAPTNILVLIHANYHHNTDSLYFMVLIATCLGSLNSCLDPLFYVMSK 378
Db 297 MYFICFAPSNNLLVVHVFYFLIKTQQRSHVYALYVALCLSLNSCIDPFPVYFVSK 351

RESULT 15

US-08-097-938-2
; Sequence 2, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097.938
; FILING DATE: 26-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22803-20006.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-097-938-2

Query Match 27.1%; Score 579.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;
Qy 95 VNNATIGYLRSSLSTQVIPAIYILLFVVGPSNIVTLWKLSLRT-KSISLVIPTHTNLIA 153
Db 59 IDEFSASILTGLTVPVVIIVFVIGLPSNGMALWIFLFTKXKHPAVIYMANLALA 118
Qy 154 DLLFCVTLFPKIAHYHLNGNNVFEVCMCRITTVFYGNMYCAIILITCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLHGNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYVWVNP 178
Qy 214 FTYQKLPRKPSFLMCGIWMVFLYMLPFVILKQEVHLVHSEITTTCHDVVDACESPSF 273
Db 179 MGHPR-KKANIYAVGSLALWLLIFLVTIPLYVMKOTIYPALNITTTCHDVLP-EVLVGD 236
Qy 274 RFFYFVSLAAGFLPIPFVIIIFCYTTLIHLKLS-----KDRILWLGVIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMKILTRSSAMDEHSEKQRRA-----IRLIITVLA 292
Qy 324 IFTICFAPTNILVLIHANYHHNTDSLYFMVLIATCLGSLNSCLDPLFYVMSK 378
Db 293 MYFICFAPSNNLLVVHVFYFLIKTQQRSHVYALYVALCLSLNSCIDPFPVYFVSK 347

RESULT 16
US-08-476-000-2
; Sequence 2, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-476-000-2

Query Match 27.1%; Score 579.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;
Qy 95 VNNATIGYLRSSLSTQVIPAIYILLFVVGPSNIVTLWKLSLRT-KSISLVIPTHTNLIA 153
Db 59 IDEFSASILTGLTVPVVIIVFVIGLPSNGMALWIFLFTKXKHPAVIYMANLALA 118
Qy 154 DLLFCVTLFPKIAHYHLNGNNVFEVCMCRITTVFYGNMYCAIILITCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLHGNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYVWVNP 178
Qy 214 FTYQKLPRKPSFLMCGIWMVFLYMLPFVILKQEVHLVHSEITTTCHDVVDACESPSF 273
Db 179 MGHPR-KKANIYAVGSLALWLLIFLVTIPLYVMKOTIYPALNITTTCHDVLP-EVLVGD 236
Qy 274 RFFYFVSLAAGFLPIPFVIIIFCYTTLIHLKLS-----KDRILWLGVIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMKILTRSSAMDEHSEKQRRA-----IRLIITVLA 292
Qy 324 IFTICFAPTNILVLIHANYHHNTDSLYFMVLIATCLGSLNSCLDPLFYVMSK 378
Db 293 MYFICFAPSNNLLVVHVFYFLIKTQQRSHVYALYVALCLSLNSCIDPFPVYFVSK 347

RESULT 17
US-08-472-840-2
; Sequence 2, Application US/08472840

```
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390.301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELEPHONE: (202) 887-0763
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-840-2

Query Match 27.1%; Score 579.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTKLTTFVLPVVIIVFVIGLPSNGMALWIFLFRTKKHPAVIYMANLALA 118

Qy 154 DLLFCVTLPPKIAVHLGNWNVFGEVCMCRITTVFVYGNMYCAIILTCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLGNWNVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRVWVIVNP 178

Qy 214 FTYOKLPKRSFSLMCGIWMVWVFLYMLPFVILKQYHLVHSEITTCDDVVDACESPSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMQTIYIPALNITTCDDVLP-EVLVGD 236

Qy 274 RFYFVSLAPFGFLIPFVIIIFCYTTLIHKLKS-----KDRWLGYIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAMDEHSENKRQRA----IRLIITVLA 292

Qy 324 IFTICFAPTNILVIHHANYHYHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 293 MYFICFAPSNNLLVHVHFLIKTORQSHVYALYVALCLSTLNSCIDPFVYFVFSK 347

RESULT 18
US-08-476-976-2
; Sequence 2, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390.301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-840-2

Query Match 27.1%; Score 579.5; DB 1; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTKLTTFVLPVVIIVFVIGLPSNGMALWIFLFRTKKHPAVIYMANLALA 118

Qy 154 DLLFCVTLPPKIAVHLGNWNVFGEVCMCRITTVFVYGNMYCAIILTCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLGNWNVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRVWVIVNP 178

Qy 214 FTYOKLPKRSFSLMCGIWMVWVFLYMLPFVILKQYHLVHSEITTCDDVVDACESPSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMQTIYIPALNITTCDDVLP-EVLVGD 236

Qy 274 RFYFVSLAPFGFLIPFVIIIFCYTTLIHKLKS-----KDRWLGYIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAMDEHSENKRQRA----IRLIITVLA 292

Qy 324 IFTICFAPTNILVIHHANYHYHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 293 MYFICFAPSNNLLVHVHFLIKTORQSHVYALYVALCLSTLNSCIDPFVYFVFSK 347

RESULT 18
US-08-476-976-2
; Sequence 2, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476.976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-2
```

```
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476.976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-2

Query Match 27.1%; Score 579.5; DB 2; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
Db 59 IDEFSASILTKLTTFVLPVVIIVFVIGLPSNGMALWIFLFRTKKHPAVIYMANLALA 118

Qy 154 DLLFCVTLPPKIAVHLGNWNVFGEVCMCRITTVFVYGNMYCAIILTCMGINRYLATAHP 213
Db 119 DLLSVIWFPLKISYHLGNWNVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRVWVIVNP 178

Qy 214 FTYOKLPKRSFSLMCGIWMVWVFLYMLPFVILKQYHLVHSEITTCDDVVDACESPSF 273
Db 179 MGHPR-KKANIAGVSLAIWLLIFLVTIPLYVMQTIYIPALNITTCDDVLP-EVLVGD 236

Qy 274 RFYFVSLAPFGFLIPFVIIIFCYTTLIHKLKS-----KDRWLGYIKAVLLILV 323
Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMIKTLRSSAMDEHSENKRQRA----IRLIITVLA 292

Qy 324 IFTICFAPTNILVIHHANYHYHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
Db 293 MYFICFAPSNNLLVHVHFLIKTORQSHVYALYVALCLSTLNSCIDPFVYFVFSK 347

RESULT 19
US-08-474-410-2
; Sequence 2, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476.976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 395 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-2
```

NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,410
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-410-2

Query Match 27.1%; Score 579.5; DB 3; Length 395;
Best Local Similarity 37.6%; Pred. No. 5.7e-37;
Matches 111; Conservative 66; Mismatches 101; Indels 17; Gaps 5;
QY 95 VNNATIGYRLSSLTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIA 153
DB 59 IDEFSASILTGLTTFVPIVYIIIVFVIGLPSNGMALWIFLFRKKKHPAVIYMANLALA 118
QY 154 DLLFCVTLPPFIYHNLGNWVFGVMCRITTVFYGNMYCAIILLTCMGINRYLATAHP 213
DB 119 DLLSVIWFPLKISYHLHGNWVYGEALCKVIGFYGNMYCSILFMTCLSVQRYWVIVNP 178
QY 214 FTYQKLPKRSFLMCGIWMVVMFLYMLPFVILKQEHVHSEITTHDVVVDACESPSF 273
DB 179 MCHPR-KKANTAVGSLALWILLFLVTIPLYVMQTIVIPALNITTCHDVLP-EVLVGD 236
QY 274 RPYFVSIAFGFLIPFVIIIFCYTTLIHLKS-----KDRIWLGVIKAVLLILV 323
DB 237 MFNYFLSLAIGVFLPALLTASAYVLMKTLRSSAMDSHSENKQRA----IRLIITVLA 292
QY 324 IFTICFAPTNILITHANYHHNTDSLYFMVLYALCLGSLNSCLDPFLYFVMSK 378
DB 293 MYFICFAPSNNLLVVHFLIKTQKSHVYALYVALCLSLNSCLDPPVYFVSK 347

RESULT 20
US-08-97-938-4
Sequence 4, Application US/08097938
Patent No. 5629174
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,938
FILING DATE: 26-JUL-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 22803-20006.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-097-938-4

Query Match 26.4%; Score 563; DB 1; Length 398;
Best Local Similarity 33.5%; Pred. No. 1e-35;
Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;
QY 49 TLTIKSPNGQPNTFEFPFLSDIEGW---TGATTTIKACPEDSISTLHVNNTATIGYLR 105
DB 21 TLVFLSCTGTNRSSKGRSLIGKVDGTSHTVGKVTVE-----TVFSDEFSSASVLTG 72
QY 106 SLSTQVIPAIIYLLFVVGVPNSIVTLWKLSLRT-KSISLVIFHTNLAIAADLLFCVTLPPK 164
DB 73 KLITVFLPIVITVIVVVGVLPSNGMALWVFLFRKKKHPAVIYMANLALADLLSVIWFPLK 132
QY 165 IAYHLNGNWNVFGVMCRITTVFYGNMYCAIILLTCMGINRYLATAHPPTYOKLPKRSF 224
DB 133 IAYHIHGNWVYGEALCNVLIGFYGNMYCSILFMTCLSVQRYWVIVNPMGHSR-KKANI 191
QY 225 SLIMCGIWMVVMFLYMLPFVILKQEHVHSEITTHDVVVDACESPSF-----RFYFVS 280
DB 192 AIGISLAIMLLILLVPIPLYVWKOTIFIPALNITTCHDVL-----PEQLVGMFNYFLS 246
QY 281 LAFFGFLIPFVIIIFCYTTLIHLKS-----KDRIWLGVIKAVLLILVIFTICFAPTNI 334
DB 247 LAIGVFLPFAFLTASAYVLMIRMLRSAMDENSEKKRKAIKLIVTVLAMYLCFTPSNL 306
QY 335 ILVHHANYHHNTDSLYFMVLYALCLGSLNSCLDPFLYFVMS 377
DB 307 LLVVHFLIKSQSHVYALYVALCLSLNSCLDPPVYFVS 349

RESULT 21
US-08-476-000-4
Sequence 4, Application US/08476000
Patent No. 5716789
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,000
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-840-4

Query Match 26.4%; Score 563; DB 1; Length 398;
Best Local Similarity 33.5%; Pred. No. 1e-35;
Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;

QY 49 TLTIKSPNGQNTPEFPFLSDIEGW---TGATTTIKACPCDSISTLHVNATIGVLR 105
DB 21 TLVFLSCTGNRSKGRSLGKVDGTSHTVKGVTVE-----TVFSVDEFSASVLTG 72
QY 106 SLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KSILVIFHTNLAIADLLFCVTLPPK 164
DB 73 KLTVFLPIVYIVFVVGSPNGMALVFLFRKKHPAVIYMANLADLLSVIWFPLK 132
QY 165 IAYHLGNWVGEVGMCRITTVVYGNMYCAILITCMGINRYLATAHPTFYQKLPKRSF 224
DB 133 IAYHIGNNMYGEALCNVLIGFFYGNMYCSILEFMTCLSVQRYVWVIVNPMGHSR-KKANI 191
QY 225 SLLMCGIIVWVFLYMLPFVILKQYHLVHSEITTCDDVDACESPSSF-----RFYFVVS 280
DB 192 AIGISLAIWLILVITPLYVVKQTFIPALNITTCDDVL-----PEQLLVGDMFNFLS 246
QY 281 LAFFGFLIPFVIIIFCYTTLIHLKLS-----KDRWLGVYKAVLLILVIFTICFAPTNI 334
DB 247 LAIGVFLFPAFLTASAYVLMRLRSSAMDENSEKKRAIKLIVTVLAMYLCFTPSNL 306
QY 335 ILVIHANYYYHNTDSLYFMYLIACLGSLNSCLDPFLYFVMS 377
DB 307 LLVWHYFLIKSQGQSHVYALYVALCLSTLNSCIDPEVYFVVS 349

RESULT 22
US-08-472-840-4
Sequence 4, Application US/08472840
Patent No. 5763575
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,840
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-840-4

Query Match 26.4%; Score 563; DB 1; Length 398;
Best Local Similarity 33.5%; Pred. No. 1e-35;
Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;

QY 49 TLTIKSPNGQNTPEFPFLSDIEGW---TGATTTIKACPCDSISTLHVNATIGVLR 105
DB 21 TLVFLSCTGNRSKGRSLGKVDGTSHTVKGVTVE-----TVFSVDEFSASVLTG 72
QY 106 SLSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KSILVIFHTNLAIADLLFCVTLPPK 164
DB 73 KLTVFLPIVYIVFVVGSPNGMALVFLFRKKHPAVIYMANLADLLSVIWFPLK 132
QY 165 IAYHLGNWVGEVGMCRITTVVYGNMYCAILITCMGINRYLATAHPTFYQKLPKRSF 224
DB 133 IAYHIGNNMYGEALCNVLIGFFYGNMYCSILEFMTCLSVQRYVWVIVNPMGHSR-KKANI 191
QY 225 SLLMCGIIVWVFLYMLPFVILKQYHLVHSEITTCDDVDACESPSSF-----RFYFVVS 280
DB 192 AIGISLAIWLILVITPLYVVKQTFIPALNITTCDDVL-----PEQLLVGDMFNFLS 246
QY 281 LAFFGFLIPFVIIIFCYTTLIHLKLS-----KDRWLGVYKAVLLILVIFTICFAPTNI 334
DB 247 LAIGVFLFPAFLTASAYVLMRLRSSAMDENSEKKRAIKLIVTVLAMYLCFTPSNL 306
QY 335 ILVIHANYYYHNTDSLYFMYLIACLGSLNSCLDPFLYFVMS 377
DB 307 LLVWHYFLIKSQGQSHVYALYVALCLSTLNSCIDPEVYFVVS 349

RESULT 23
US-08-476-976-4
Sequence 4, Application US/08476976
Patent No. 5874400
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-0763
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-4

Query Match 26.4%; Score 563; DB 2; Length 398;

Best Local Similarity 33.5%; Pred. No. 1e-35;
Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;

```
QY 49 TLTKSFNGGQNTPEEFPLSDIEGW---TGATTIKAECPEDSISTLHVNNATIGVLR 105
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 TLVFLSCTGTRSSKGRSLIKGVDTGSHVTKGVTV-----TVFSVDEFSASVLTG 72

QY 106 SLSTQVPIAYILLFVVGVPNSIVTLWKLRLT-KSISLVIFHTNLAIADLLFCVTLFPK 164
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 73 KLTVFLPIVYTVFVVGGLPSNGMALWFLRTRKKHPPAVIYMANLADLLSVIWPFLK 132

QY 165 IAYHLGNWVGEVWMCRTTIVFYGNMYCAILLTCMGINRYLATAHPTYOKLPKRSF 224
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 IAYHIGNNWVGEALCNVLIGFFYGNMYCSILFMTCLSVQRYVWVNPNGHSR-KKANI 191

QY 225 SLLMCGIVWVWVFLYMLPFVILKQEHVHSEITTCCHVDVDACESPSF-----RFYFVS 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 AIGISLAIWLLLVITPLVYVVKQTIFIPALNITTCCHDL-----PEQLLVGDMFNFLS 246

QY 281 LAFFGFLIPFVIIIFYCTTLIHKLK-----KDRIMLGYIKAVLLILVITTCFAPTNI 334
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 LAIGVFLPAFLTASAVVLMIRLSSAMDENSEKKRKAIKLIVTVLWVLYICFTPSNL 306

QY 335 ILVTHHANYYYHTDSLFWYLLIALCLGSLNSCLDPLFYFVS 377
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 LLVHYFLIKSQGQSHVYALVALCLSTLNSCIDPFFVYFVS 349
```

RESULT 24

US-08-474-410-4
; Sequence 4, Application US/08/474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER

; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-410-4

Query Match 26.4%; Score 563; DB 3; Length 398;

Best Local Similarity 33.5%; Pred. No. 1e-35;
Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;

```
QY 49 TLTKSFNGGQNTPEEFPLSDIEGW---TGATTIKAECPEDSISTLHVNNATIGVLR 105
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 TLVFLSCTGTRSSKGRSLIKGVDTGSHVTKGVTV-----TVFSVDEFSASVLTG 72

QY 106 SLSTQVPIAYILLFVVGVPNSIVTLWKLRLT-KSISLVIFHTNLAIADLLFCVTLFPK 164
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 73 KLTVFLPIVYTVFVVGGLPSNGMALWFLRTRKKHPPAVIYMANLADLLSVIWPFLK 132

QY 165 IAYHLGNWVGEVWMCRTTIVFYGNMYCAILLTCMGINRYLATAHPTYOKLPKRSF 224
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 IAYHIGNNWVGEALCNVLIGFFYGNMYCSILFMTCLSVQRYVWVNPNGHSR-KKANI 191

QY 225 SLLMCGIVWVWVFLYMLPFVILKQEHVHSEITTCCHVDVDACESPSF-----RFYFVS 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 AIGISLAIWLLLVITPLVYVVKQTIFIPALNITTCCHDL-----PEQLLVGDMFNFLS 246

QY 281 LAFFGFLIPFVIIIFYCTTLIHKLK-----KDRIMLGYIKAVLLILVITTCFAPTNI 334
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 247 LAIGVFLPAFLTASAVVLMIRLSSAMDENSEKKRKAIKLIVTVLWVLYICFTPSNL 306

QY 335 ILVTHHANYYYHTDSLFWYLLIALCLGSLNSCLDPLFYFVS 377
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 LLVHYFLIKSQGQSHVYALVALCLSTLNSCIDPFFVYFVS 349
```

RESULT 25

US-08-486-673B-4
; Sequence 4, Application US/08/486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07

```

; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-486-673B-4

Query Match      26.4%; Score 563; DB 4; Length 398;
Best Local Similarity 33.5%; Pred. No. 1.4e-35;
Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;

Qy 49 TLTKSFNGGPNTEEFPSLDIEG---TGATTIKAECPEDSISTLHVNNATIGYLR 105
Db 21 TLVFLSTGNTNRSSKGRSLGKVDGTSHTVKGVTVE-----TVFSVDEFSASVLTG 72

Qy 106 SLSTQVIPAIIYLLFVVGPSNIVTLWKLSLRT-KSISLVIFHTNLAIADLLFCVTLPPK 164
Db 73 KLTTFELPIVTVIVFVGLPSNGMALWFLPRTKKHPAVIYMANLADLLSVIWFPLK 132

Qy 165 IAYHLGNWVGEVMCRTITTVFYGNMYCAIILTCMGINRYLATAPHTYQKLPKRSF 224
Db 133 IAYHIGNWNYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSR-KKANI 191

Qy 225 SLLMCGIYVWVFLYMLPFVILKQYHLVHSEITTCDDVVDACSPSSF----RFYVFS 280
Db 192 AIGSLALWLLILVTLPIYVVKQTIIPALNITTCDDVL-----PQLLVGDMFNFLS 246

Qy 281 LAFFGFLIPFVIIIFCYTTLIHKLK-----KDRILWGLYIKAVLLILVIPTICFAPTNI 334
Db 247 LAIGVFLPAPLTASAYVLMRLMRSSAMDENSEKKRAIKLIVTVLWMLYLCFTSNL 306

Qy 335 ILVTHANYYYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 307 LLVVHYFLIKSQGSHVYALVALCLSLTNSCIDPFVYVFS 349

RESULT 26
US-08-486-673B-6
; Sequence 6, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; APPLICANT: Scarborough, Robert M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Protein sequence from cDNA
US-08-486-673B-6

Query Match      26.3%; Score 561.5; DB 4; Length 398;
Best Local Similarity 36.7%; Pred. No. 1.4e-35;
Matches 108; Conservative 63; Mismatches 106; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVVGPSNIVTLWKLSLRT-KSISLVIFHTNLAI 153
Db 62 VDEFSASVLTGKLTIVFLPIVYIIVFVGLPSNGMALWFLRTKKKHPAVIYMANLALA 121

Qy 154 DLLFCVTLPPFKIAYHLGNWVGEVMCRTITTVFYGNMYCAIILTCMGINRYLATAPH 213
Db 122 DLLSVIWFPLPKIAYHIGNWNYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVN 181

Qy 214 FTYQKLPKRSFSLMCGIYVWVFLYMLPFVILKQYHLVHSEITTCDDVVDACSPSSF 273
Db 182 MGHSR-KKANIAIGISLALWLLILVTLPIYVVKQTIIPALNITTCDDVL-----PEQL 235

Qy 274 ---REVYFVSFLAFFGLIPFVIIIFCYTTLIHKLK-----KDRILWGLYIKAVLLILV 323
Db 236 LVGDMFNFLSALIGVFLPAPLTASAYVLMRLMRSSAMDENSEKKRAIKLIVTVLA 295

Qy 324 IFTICFAPTNIILVTHANYYYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 296 MYLICFIPSNLLVHVHYFLIKSQGSHVYALVALCLSLTNSCIDPFVYVFS 349

RESULT 27
US-08-097-938-6
; Sequence 6, Application US/08097938
; Patent No. 5629174
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,938
; FILING DATE: 26-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22803-20006.00
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-097-938-6

Query Match      26.2%; Score 560.5; DB 1; Length 398;
Best Local Similarity 36.7%; Pred. No. 1.6e-35;
Matches 108; Conservative 63; Mismatches 106; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVVGPSNIVTLWKLSLRT-KSISLVIFHTNLAI 153
Db 62 VDEFSASVLTGKLTIVFLPIVYIIVFVGLPSNGMALWFLRTKKKHPAVIYMANLALA 121

Qy 154 DLLFCVTLPPFKIAYHLGNWVGEVMCRTITTVFYGNMYCAIILTCMGINRYLATAPH 213
Db 122 DLLSVIWFPLPKIAYHIGNWNYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVN 181

Qy 214 FTYQKLPKRSFSLMCGIYVWVFLYMLPFVILKQYHLVHSEITTCDDVVDACSPSSF 273
Db 182 MGHSR-KKANIAIGISLALWLLILVTLPIYVVKQTIIPALNITTCDDVL-----PEQL 235
```

Db 182 MHSR-KKANIAIGISLAIMLLVLTPIVYVVKQTIFIPALNITTTCHDVL-----PEQV 235
Qy 274 ----RFFYVSLAFGLFIPVILFCVTTLIHKLK-----KRIWLGVIKAVLILV 323
Db 236 LVGDMFNFLSLAIGVFLFPAFLTASAYVLMIRLSSAMDNSEKKRAIKLIVTLA 295
Qy 324 IFTICFAPTNILVIHHANYHNTDSLYFMYLIALCLGSLNSCLDPLFYFVMS 377
Db 296 MYLICFIPSNLLLVVHFLIKSQGSHVYALIVLCLSTLNSCIDPFVYFVS 349

RESULT 28
US-08-476-000-6
; Sequence 6, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-476-000-6

Query Match 26.2%; Score 560.5; DB 1; Length 398;
Best Local Similarity 36.7%; Pred. No. 1.6e-35;
Matches 108; Conservative 63; Mismatches 106; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSQTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KSISLVIFHTNLAIA 153
Db 62 VDEFSASVLTGKLTITVFLPIVYIIVFVGLPSNGMALWVFLFRKKHPAIVYMANLALA 121
Qy 154 DLLFCVTLPPKIAVHLNNGNWFGEVMCRITTVVFGNMYCAIILITCMGINRYLATAHP 213
Db 122 DLLSVIWFPLKIAVHIGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 181
Qy 214 FTYOKLPKRSFLMCGIWMVFLYMLPFVILKQEHVHSEITTTCHDVVDACESPSS 273
Db 182 MHSR-KKANIAIGISLAIMLLVLTPIVYVVKQTIFIPALNITTTCHDVL-----PEQV 235
Qy 274 ----RFFYVSLAFGLFIPVILFCVTTLIHKLK-----KRIWLGVIKAVLILV 323
Db 296 MYLICFIPSNLLLVVHFLIKSQGSHVYALIVLCLSTLNSCIDPFVYFVS 349

Db 236 LVGDMFNFLSLAIGVFLFPAFLTASAYVLMIRLSSAMDNSEKKRAIKLIVTLA 295
Qy 324 IFTICFAPTNILVIHHANYHNTDSLYFMYLIALCLGSLNSCLDPLFYFVMS 377
Db 296 MYLICFIPSNLLLVVHFLIKSQGSHVYALIVLCLSTLNSCIDPFVYFVS 349

RESULT 29
US-08-472-840-6
; Sequence 6, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-472-840-6

Query Match 26.2%; Score 560.5; DB 1; Length 398;
Best Local Similarity 36.7%; Pred. No. 1.6e-35;
Matches 108; Conservative 63; Mismatches 106; Indels 17; Gaps 5;

Qy 95 VNNATIGYLRSSLSQTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KSISLVIFHTNLAIA 153
Db 62 VDEFSASVLTGKLTITVFLPIVYIIVFVGLPSNGMALWVFLFRKKHPAIVYMANLALA 121
Qy 154 DLLFCVTLPPKIAVHLNNGNWFGEVMCRITTVVFGNMYCAIILITCMGINRYLATAHP 213
Db 122 DLLSVIWFPLKIAVHIGNNWIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 181
Qy 214 FTYOKLPKRSFLMCGIWMVFLYMLPFVILKQEHVHSEITTTCHDVVDACESPSS 273
Db 182 MHSR-KKANIAIGISLAIMLLVLTPIVYVVKQTIFIPALNITTTCHDVL-----PEQV 235
Qy 274 ----RFFYVSLAFGLFIPVILFCVTTLIHKLK-----KRIWLGVIKAVLILV 323
Db 236 LVGDMFNFLSLAIGVFLFPAFLTASAYVLMIRLSSAMDNSEKKRAIKLIVTLA 295

QY 324 IFTICFAPTNIIIVHHANYHNTDSLYFMVLIACLSGNSCLDPFLYFVMS 377
 Db 296 MYLICFIPSNLLLVVHFLIKSQSQSHVYALIVLCLSTLNSCIDPFFVYFVS 349

RESULT 30
 US-08-476-976-6
 ; Sequence 6, Application US/08476976
 ; Patent No. 5874400
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/476,976
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 398 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-476-976-6

Query Match 26.2%; Score 560.5; DB 2; Length 398;
 Best Local Similarity 36.7%; Pred. No. 1.6e-35;
 Matches 108; Conservative 63; Mismatches 106; Indels 17; Gaps 5;
 QY 95 VNNATIGYLRSSISTQVIPAIIYLLFVVGVPSPNIVLWKLRLT-KSISIVIFHTNLATA 153
 Db 62 VDEFSASVLTGKTTVPFIYIIVFVVGLPNGMALWFLFRTKKKHPAVIYMANLALA 121
 QY 154 DLLFCVTLTPFKIAYHLNGNWNVGEVWCRITTVVFGNMYCAILILTCMGINRYLATAHP 213
 Db 122 DLLSVIWFPLKIAHYHGNWYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 181
 QY 214 FTYOKLPKRSFSLMCGIWMVFLMPLFVILKQYHLVHSEITTCDDVVDACESPSF 273
 Db 182 MGHRSR-KKANAIGISLAIWLLIIVPLVYVVKQTFIPALNITTCDDVL-----PEQV 235
 QY 274 ----RFYVYVSLAPFGFLIPFVIIIFCYTTLIHKLK-----KDRWLGYIKAVLLILV 323
 Db 236 LVGDMFNFLSLAIGVLFPAFLTASAYVLMIRMLRSSAMENSEKKRRAIKLIVTLA 295
 QY 324 IFTICFAPTNIIIVHHANYHNTDSLYFMVLIACLSGNSCLDPFLYFVMS 377
 Db 296 MYLICFIPSNLLLVVHFLIKSQSQSHVYALIVLCLSTLNSCIDPFFVYFVS 349

RESULT 32

US-08-474-410-6
 ; Sequence 6, Application US/08474410
 ; Patent No. 6043212
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,410
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 398 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-474-410-6

Query Match 26.2%; Score 560.5; DB 3; Length 398;
 Best Local Similarity 36.7%; Pred. No. 1.6e-35;
 Matches 108; Conservative 63; Mismatches 106; Indels 17; Gaps 5;
 QY 95 VNNATIGYLRSSISTQVIPAIIYLLFVVGVPSPNIVLWKLRLT-KSISIVIFHTNLATA 153
 Db 62 VDEFSASVLTGKTTVPFIYIIVFVVGLPNGMALWFLFRTKKKHPAVIYMANLALA 121
 QY 154 DLLFCVTLTPFKIAYHLNGNWNVGEVWCRITTVVFGNMYCAILILTCMGINRYLATAHP 213
 Db 122 DLLSVIWFPLKIAHYHGNWYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 181
 QY 214 FTYOKLPKRSFSLMCGIWMVFLMPLFVILKQYHLVHSEITTCDDVVDACESPSF 273
 Db 182 MGHRSR-KKANAIGISLAIWLLIIVPLVYVVKQTFIPALNITTCDDVL-----PEQV 235
 QY 274 ----RFYVYVSLAPFGFLIPFVIIIFCYTTLIHKLK-----KDRWLGYIKAVLLILV 323
 Db 236 LVGDMFNFLSLAIGVLFPAFLTASAYVLMIRMLRSSAMENSEKKRRAIKLIVTLA 295
 QY 324 IFTICFAPTNIIIVHHANYHNTDSLYFMVLIACLSGNSCLDPFLYFVMS 377
 Db 296 MYLICFIPSNLLLVVHFLIKSQSQSHVYALIVLCLSTLNSCIDPFFVYFVS 349

US-08-486-673B-63
; Sequence 63, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Sundelin, Johan
; APPLICANT: Scarborough, Robert M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 63
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-486-673B-63

Query Match 26.2%; Score 559; DB 4; Length 397;
Best Local Similarity 32.4%; Pred. No. 2.1e-35;
Matches 119; Conservative 72; Mismatches 138; Indels 38; Gaps 8;
Qy 22 LVAAGLLFLPVTVCQSGINVDNSAKPTLTIKSFNGSPQNTFEFPPLSDIEGWTGATTI 81
Db 9 LLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGRGVTV 55
Qy 82 KAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPSTNITLWKLRLT-KS 140
Db 56 E-----TVFSVDEFSASVLTKGTTVPFLPIVTTIVFVGLPSNGMALWVFLFRYKK 107
Qy 141 ISLVIFHTNLAIADLLFCVTLPPFKIAYHLNGNNWVFGVMCRITTVVFGNMYCAILIT 200
Db 108 HPVIVYMANLADLLSVIWFPLKIAHYHGNWNIYGEALCNVLIGFFYGNMYCSILFMT 167
Qy 201 CMGINRYLATAPTYQKLPKRSFSLMCGIWMVFLYMLPFVILKQEHVHSEITTC 260
Db 168 CLSVQRYWVIVNPMGHSR-KKANAIGISLAIWLLTLTLVTPLYVVKQTIFIPALNITTC 226
Qy 261 HDVVDAESPSSF---RFYFVSFAFFGLPIPVIIIFCVTTLIHLKS-----KDIR 310
Db 227 HDVL-----PQLLVGDMFNFLSLAIGVFLFPALFASAVLMIRLRSAMDENSEKK 281
Qy 311 WLGIKAVLLILVITTCFAPTNIIIVHANYHHNTDSLYFMYLIALCLGSLNSCLDP 370
Db 282 RKRAIKLIVTLGMYLICFTSPNLLLVVHYFLIKSQGSHVYALYIVALCLSLTNSCIDP 341
Qy 371 FLYFVMS 377
Db 342 VYFVVS 348

RESULT 33
US-08-742-440A-8
; Sequence 8, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicvic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,440A
FILING DATE: 30-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: UCAL/060PAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3400
TELEFAX: 650 327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 394 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-440A-8

Query Match 25.9%; Score 553.5; DB 2; Length 394;
Best Local Similarity 32.2%; Pred. No. 5.5e-35;
Matches 118; Conservative 73; Mismatches 136; Indels 39; Gaps 8;
Qy 22 LVAAGLLFLPVTVCQSGINVDNSAKPTLTIKSFNGSPQNTFEFPPLSDIEGWTGATTI 81
Db 9 LLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGRGVTV 55
Qy 82 KAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPSTNITLWKLRLT-KS 140
Db 56 E-----TVFSVDEFSASVLTKGTTVPFLPIVTTIVFVGLPSNGMALWVFLFRYKK 107
Qy 141 ISLVIFHTNLAIADLLFCVTLPPFKIAYHLNGNNWVFGVMCRITTVVFGNMYCAILIT 200
Db 108 HPVIVYMANLADLLSVIWFPLKIAHYHGNWNIYGEALCNVLIGFFYGNMYCSILFMT 167
Qy 201 CMGINRYLATAPTYQKLPKRSFSLMCGIWMVFLYMLPFVILKQEHVHSEITTC 260
Db 168 CLSVQRYWVIVNPMGHSR-KKANAIGISLAIWLLTLTLVTPLYVVKQTIFIPALNITTC 226
Qy 261 HDVVDAESPSSF---RFYFVSFAFFGLPIPVIIIFCVTTLIHLKS-----KDIR 311
Db 227 HDVLPEQLLVGDP-----FLSLAIGVFLFPALFASAVLMIRLRSAMDENSEKK 279
Qy 312 LGYIKAVLLILVITTCFAPTNIIIVHANYHHNTDSLYFMYLIALCLGSLNSCLDP 371
Db 280 KRAIKLIVTLGMYLICFTSPNLLLVVHYFLIKSQGSHVYALYIVALCLSLTNSCIDP 339
Qy 372 LYFVMS 377
Db 340 VYFVVS 345

RESULT 34
US-08-476-000-63
; Sequence 63, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,000
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-000-63

Query Match 25.8%; Score 551; DB 1; Length 397;
Best Local Similarity 32.2%; Pred. No. 8.5e-35;
Matches 118; Conservative 72; Mismatches 139; Indels 38; Gaps 8;
QY 22 LVAAGLLFLPVTVCQSGINVDNSAKPTLTIKSFNGGPQNTFEEPLSDIEGWTGATTI 81
DB 9 LLGNAILLAAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
QY 82 KAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KS 140
DB 56 E-----TVFSVDEFSASVLAGKLTTPVITVFAVGLPSNGMALWVFLFRKK 107
QY 141 ISLVIFHTNLAIADLLFCVTLFPKIAIYHLNNGNWFGEVNCRTITTVFYGNMYCAIILT 200
DB 108 HPVIVYMANLALADLLSVIWFPLKIAIYHNGNWIYGEALCNVLIGFFYRNMYSILFMT 167
QY 201 CMGINRYLATAHPFTYQKLPKRSFLLMCGIWMVFLYMLPFVILKQYHLVHSEITTC 260
DB 168 CLSVQRVWVIVNPMGHSR-KKANIAIGISLAIWLLTLVTIPLYVVKQTFIFALNITTC 226
QY 261 HDVVDAACESPSF-----RFYFVSLAFPGFLIPFVIIIFCYTTLIHKLK-----KDRI 310
DB 227 HDVL-----PEQLLVGDMFNFLSLAIGVFLPAPFLTASAYVLMIRLSSAMDENSEK 281
QY 311 WLGIYKAVLLILVIFTCFAPTNIIIVHANNYYHNTDLSLYFMYIALCLGSLNSCLDP 370
DB 282 RKRAKILVTVLGMYLICFTPSNLLLVVHYFLIKSQSQSHVYALYIIVALCLSTLNSCIDP 341
QY 371 FLYFVMS 377
DB 342 FVYFVS 348

RESULT 35
US-08-472-840-63
Sequence 63 Application US/08472840
Patent No. 5763575
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN

APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,840
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-840-63

Query Match 25.8%; Score 551; DB 1; Length 397;
Best Local Similarity 32.2%; Pred. No. 8.5e-35;
Matches 118; Conservative 72; Mismatches 139; Indels 38; Gaps 8;
QY 22 LVAAGLLFLPVTVCQSGINVDNSAKPTLTIKSFNGGPQNTFEEPLSDIEGWTGATTI 81
DB 9 LLGNAILLAAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
QY 82 KAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KS 140
DB 56 E-----TVFSVDEFSASVLAGKLTTPVITVFAVGLPSNGMALWVFLFRKK 107
QY 141 ISLVIFHTNLAIADLLFCVTLFPKIAIYHLNNGNWFGEVNCRTITTVFYGNMYCAIILT 200
DB 108 HPVIVYMANLALADLLSVIWFPLKIAIYHNGNWIYGEALCNVLIGFFYRNMYSILFMT 167
QY 201 CMGINRYLATAHPFTYQKLPKRSFLLMCGIWMVFLYMLPFVILKQYHLVHSEITTC 260
DB 168 CLSVQRVWVIVNPMGHSR-KKANIAIGISLAIWLLTLVTIPLYVVKQTFIFALNITTC 226
QY 261 HDVVDAACESPSF-----RFYFVSLAFPGFLIPFVIIIFCYTTLIHKLK-----KDRI 310
DB 227 HDVL-----PEQLLVGDMFNFLSLAIGVFLPAPFLTASAYVLMIRLSSAMDENSEK 281
QY 311 WLGIYKAVLLILVIFTCFAPTNIIIVHANNYYHNTDLSLYFMYIALCLGSLNSCLDP 370
DB 282 RKRAKILVTVLGMYLICFTPSNLLLVVHYFLIKSQSQSHVYALYIIVALCLSTLNSCIDP 341
QY 371 FLYFVMS 377
DB 342 FVYFVS 348

RESULT 36

US-08-476-976-63
; Sequence 63, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-976-63

Query Match 25.8%; Score 551; DB 2; Length 397;
Best Local Similarity 32.2%; Pred. No. 8.5e-35;
Matches 118; Conservative 72; Mismatches 139; Indels 38; Gaps 8;

Qy 22 LVAAGLLFLPVTVCQSGINVDNSAKPTLTIKSFNGGPQNTFEEFPLSDIEGWTGATTI 81
Db 9 LLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
Qy 82 KAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KS 140
Db 56 E-----TVFSVDEFSASVLAKLTTFVLPVITYTVFAVGLPSNGMALWVFLFRKK 107
Qy 141 ISLVIFHTNLAIADLLFCVTLPEKIAVHLNNGWVGEVMCRTITTVFYGNMYCAIILT 200
Db 108 HPVIVMANLADLLSVIWFDPKIAVHNGNWIYGEALCNVLIQFFYRNMYSILFMT 167
Qy 201 CMGINRYLATAPFTYQKLPKRSFSLMCGIWMVWVFLYMLPFVILKQEHVHSEITTC 260
Db 168 CLSVQRWVIVNPMGHR-KKANIAGISLAIWLLTLVTIPLVYVVKQTFIPALNITTC 226
Qy 261 HDVVDAESPSSP-----RFYFVSLAFPGFLIPFVIIIFCYTTLIHLKLS-----KDRI 310
Db 311 WLGVIKAVLLIIVITFCFAPTNILVIHANNYYHNTDSLYFMYLALCLGSLNSCLDP 370
Db 282 RKRAIKLIVTGLMGLICFTPSNLLVHVHFLIKSQSQSHVYALYIIVALCLTSLNSCIDP 341
Qy 371 FLYFVMS 377

Db 342 FVYFVUS 348
RESULT 37
US-08-474-410-63
; Sequence 63, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-474-410-63

Query Match 25.8%; Score 551; DB 3; Length 397;
Best Local Similarity 32.2%; Pred. No. 8.5e-35;
Matches 118; Conservative 72; Mismatches 139; Indels 38; Gaps 8;

Qy 22 LVAAGLLFLPVTVCQSGINVDNSAKPTLTIKSFNGGPQNTFEEFPLSDIEGWTGATTI 81
Db 9 LLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
Qy 82 KAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLRLT-KS 140
Db 56 E-----TVFSVDEFSASVLAKLTTFVLPVITYTVFAVGLPSNGMALWVFLFRKK 107
Qy 141 ISLVIFHTNLAIADLLFCVTLPEKIAVHLNNGWVGEVMCRTITTVFYGNMYCAIILT 200
Db 108 HPVIVMANLADLLSVIWFDPKIAVHNGNWIYGEALCNVLIQFFYRNMYSILFMT 167
Qy 201 CMGINRYLATAPFTYQKLPKRSFSLMCGIWMVWVFLYMLPFVILKQEHVHSEITTC 260
Db 168 CLSVQRWVIVNPMGHR-KKANIAGISLAIWLLTLVTIPLVYVVKQTFIPALNITTC 226
Qy 261 HDVVDAESPSSP-----RFYFVSLAFPGFLIPFVIIIFCYTTLIHLKLS-----KDRI 310
Db 227 HDVL-----PEQLLVGDMFNYFLSLAIGVFLPFAFLTASAYVLMIRLMRSSAMENSEK 281

QY 311 WLGVKAVLLLVLTFTICFAPTNILVHHANYHNTDLSLYFMYLIALCLGSLNSCLDP 370
 Db 282 RKRAKLIVTGLVGLYCFTPSNLLLVVHYFLIKSQGSHVYALYVALCLSTNSCIDP 341
 QY 371 FLYFVMS 377
 Db 342 FVYFVS 348

RESULT 38

US-07-657-769B-69
 ; Sequence 69, Application US/07657769B
 ; Patent No. 5256766
 ; GENERAL INFORMATION:
 ; APPLICANT: COUGHLIN, SHAUN R.
 ; TITLE OF INVENTION: RECOMBINANT RECEPTOR AND RELATED
 ; PHARMACEUTICALS
 ; NUMBER OF SEQUENCES: 69
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: IRELL & MANELLA
 ; STREET: 545 MIDDLEFIELD ROAD, SUITE 200
 ; CITY: MENLO PARK
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/657,769B
 ; FILING DATE: 19910219
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 2000-0502.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-327-7250
 ; TELEFAX: 415-327-2951
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 69:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 425 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-657-769B-69

Query Match 23.8%; Score 509; DB 1; Length 425;
 Best Local Similarity 31.2%; Pred. No. 1.5e-31;
 Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;
 QY 20 LILVAAGLLFLPVTVCOSGINVSDNSAKP-----TLTIKSF-NGGPONTFEEFPLSDI 71
 Db 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDRSFLLRNPNDKYEPFWEDEE 60
 QY 72 EGMTGAT-----TTIKACPEDSISTLHVNNATIGYLRSSLSQTQVIPAIIYILLFVVGVPNS 127
 Db 61 KNESGLTEYRLVSVINKSSPLQKLPAPFISEDAGSYLTSSWLTFLVPSVYTGVFVVSPLN 120
 QY 128 I--VTLWKLRLTKSISLVIFHTNLAIADLLFCVTLFPKIAYHLNGNNVFGVEMCRITT 185
 Db 121 IMAIVVFLKMKYK-K-PAVVYMLHLATADVLFVSVLPFKISYFSGSDWQFGSELCHFVT 179
 QY 186 VVFGNNMYCAIILTCMGINRYLATAHP---FTYQKLPKRSFSLMCGIWMVVMFLYMLP 242
 Db 180 AAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASFT---CLAIWALAIAGVWP 236
 QY 243 FVILKQEHVHSEITTCCHVDVDAKSPSFRFYFVSLAFAFGFLIPFVIIIFCYTTLIH 302
 Db 237 LVLKEQTIQVPLNITTCCHVDVDAKSPSFRFYFVSLAFAFGFLIPFVIIIFCYTTLIH 295

QY 303 KLKS-----KDRIMLVGVIKAVLL---ILVIFTICFAPTNILVHHANYHNTDLSLYFMYLIALCLGSLNSCLDP 370
 Db 296 CLSSSAVANRSKSR-----ALFLSAAVFCIFCIGFTNVLLIAHY-SFLSHTSTTE 347
 QY 350 SLYFMYLIALCLGSLNSCLDPFLYFVMS 377
 Db 348 AAYFAYLLCVCVSSISSCIDPLIYYVAS 375

RESULT 39

US-07-789-184-220
 ; Sequence 220, Application US/07789184
 ; Patent No. 5688768
 ; GENERAL INFORMATION:
 ; APPLICANT: COUGHLIN, SHAUN R.
 ; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
 ; RELATED PHARMACEUTICALS
 ; NUMBER OF SEQUENCES: 223
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 Page Mill Road
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/789,184
 ; FILING DATE: 19911107
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MURASHIGE, KATE H.
 ; REGISTRATION NUMBER: 29,959
 ; REFERENCE/DOCKET NUMBER: 22000-20502.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 813-5600
 ; TELEFAX: (415) 494-0792
 ; TELEX: 34-0154
 ; INFORMATION FOR SEQ ID NO: 220:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 425 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-789-184-220

Query Match 23.8%; Score 509; DB 1; Length 425;
 Best Local Similarity 31.2%; Pred. No. 1.5e-31;
 Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;
 QY 20 LILVAAGLLFLPVTVCOSGINVSDNSAKP-----TLTIKSF-NGGPONTFEEFPLSDI 71
 Db 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDRSFLLRNPNDKYEPFWEDEE 60
 QY 72 EGMTGAT-----TTIKACPEDSISTLHVNNATIGYLRSSLSQTQVIPAIIYILLFVVGVPNS 127
 Db 61 KNESGLTEYRLVSVINKSSPLQKLPAPFISEDAGSYLTSSWLTFLVPSVYTGVFVVSPLN 120
 QY 128 I--VTLWKLRLTKSISLVIFHTNLAIADLLFCVTLFPKIAYHLNGNNVFGVEMCRITT 185
 Db 121 IMAIVVFLKMKYK-K-PAVVYMLHLATADVLFVSVLPFKISYFSGSDWQFGSELCHFVT 179
 QY 186 VVFGNNMYCAIILTCMGINRYLATAHP---FTYQKLPKRSFSLMCGIWMVVMFLYMLP 242
 Db 180 AAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASFT---CLAIWALAIAGVWP 236
 QY 243 FVILKQEHVHSEITTCCHVDVDAKSPSFRFYFVSLAFAFGFLIPFVIIIFCYTTLIH 302

Db 237 LVLKEQTIQVGLNITTCDDVNETLEGYYAYFSAFSAVF-FFVPLIISTVCYYSIIR 295
QY 303 KLKS-----KDRIWLGVIKAVLL---ILVIFTICFAPTNILVHANYHH--NTD 349
Db 296 CLSSSAVANRSKSR-----ALFLSAAVFCIFICFGPTNVLLIAHY-SFLSHTSTTE 347
QY 350 SLYFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 40

US-08-475-263-220
; Sequence 220, Application US/08475263
; Patent No. 5759994

; GENERAL INFORMATION:

; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA

; ZIP: 20006-1812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/475,263

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: MURASHIGE, KATE H.

; REGISTRATION NUMBER: 29,959

; REFERENCE/DOCKET NUMBER: 22000-20502.03

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 887-1500

; TELEFAX: (202) 887-0763

; TELEX: 90-4030

; INFORMATION FOR SEQ ID NO: 220:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 425 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-475-263-220

Query Match 23.8%; Score 509; DB 1; Length 425;
Best Local Similarity 31.2%; Pred. No. 1.5e-31;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

QY 20 LILVAAGLLFLPVTVCOSGINVDSNAKP-----TLTIKSP-NGGPQNTFEFFPLSDI 71
Db 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDPKSLRNPNPKYEPFWEDEE 60
QY 72 EGWTGAT---TTIKACPEDSISTLHVNNATIGYLRSSLSIQVPAIYILLFVWGVPSN 127
Db 61 KNESGLTEYRLVLSINKSPLOKQLPAFISEDAGYLTSSWLTLPVPSVYTGVPVWSLPLN 120
QY 128 I--VTLKLSLRKTSISLVIFHNLAIADLLFCVTLPEFKIAYHLNGNNWVFGVMCRITT 185
Db 121 IMAIWFIKMKVKK-PAVVVYMLHLATADLVFVSVLPFKISYYPSGSDWQFGSELCRFVT 179
QY 186 VVFYGNMYCALLITCNGINRYLATAHP---FTYQKLPKRSFSLMCGIYVWVFLYMLP 242
Db 180 AAFYCNMYAIIIMTVISIDRFLLAVVYPMQSLSWRTLGRASFT---CLAIWALAIAGVVP 236

QY 243 FVILKOEYHLVHSEIITTCDDVWDACESPSSFRFYFVSLAFFGFLIPFVILIFCYTTLIH 302
Db 237 LVLKEQTIQVGLNITTCDDVNETLEGYYAYFSAFSAVF-FFVPLIISTVCYYSIIR 295
QY 303 KLKS-----KDRIWLGVIKAVLL---ILVIFTICFAPTNILVHANYHH--NTD 349
Db 296 CLSSSAVANRSKSR-----ALFLSAAVFCIFICFGPTNVLLIAHY-SFLSHTSTTE 347
QY 350 SLYFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375

Search completed: June 24, 2003, 12:07:18

Job time : 20.9779 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:52:41 ; Search time 51.9362 Seconds
(without alignments)
1044.224 Million cell updates/sec

Title: US-09-208-629F-3
Perfect score: 2136
Sequence: 1 TLTXQHPVAGSODIKMKIL.....AMARPLXPRRDIWIHAW 407

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_101002.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1940	90.8	369	19 AAW51405	Mouse protease-act
2	1425.5	66.7	374	19 AAW51406	Human protease-act
3	581.5	27.2	399	16 AAR66922	Murine Cl40 recept
4	581.5	27.2	399	17 AAW01954	Murine Cl40 recept
5	574.5	26.9	395	16 AAR66920	Murine Cl40 recept
6	572.5	26.8	395	17 AAW01952	Murine Cl40 recept
7	569	26.6	397	21 AAB35641	Human PAR-2 protei
8	563	26.4	398	16 AAR66921	Human Cl40 recepto
9	563	26.4	398	17 AAW01953	Human Cl40 recepto
10	553.5	25.9	394	19 AAW51408	Human protease-act

11	551	25.8	397	16 AAR66923	Human Cl40 recept
12	551	25.8	397	17 AAW01955	Human Cl40 recepto
13	509.5	23.9	385	20 AAY50139	Mutant human prote
14	509.5	23.9	408	20 AAY50137	FLAG epitope-tagge
15	509	23.8	425	13 AAR27240	Human thrombin rec
16	509	23.8	425	19 AAW51407	Human protease-act
17	509	23.8	425	20 AAY49570	Human thrombin rec
18	509	23.8	425	23 AAG80697	Human thrombin pro
19	509	23.8	425	23 AAE17032	Human thrombin pro
20	509	23.8	426	21 AAY45035	Human thrombin rec
21	507.5	23.8	385	20 AAY15082	Human protease-act
22	507.5	23.8	385	20 AAY50135	Human protease-act
23	507.5	23.8	385	21 AAY45036	Human protease act
24	507.5	23.8	385	22 AAB47623	Human PAR4. Homo
25	507	23.7	425	15 AAR60698	Fragment of the hu
26	506.5	23.7	385	20 AAY50138	Mutant human prote
27	502	23.5	892	18 AAW16314	Human thrombin rec
28	501	23.5	425	23 AEG35300	Human PAR1 type th
29	494	23.1	359	21 AAY71304	Human orphan G pro
30	494	23.1	359	21 AAB02838	Human G protein co
31	494	23.1	396	20 AAY15081	Mouse protease-act
32	493	23.1	359	21 AAY84815	A human G-protein
33	493	23.1	359	22 AAG78530	G-protein coupled
34	493	23.1	359	22 AAG80966	Human nPCR5 #2.
35	493	23.1	359	22 AAB62285	Human G-protein co
36	489	22.9	359	21 AAY69485	Amino acid sequenc
37	485.5	22.7	402	23 AAG35298	Human PAR1 type th
38	484.5	22.7	371	23 AAG35299	Human PAR1 type th
39	444.5	20.8	359	19 AAW69598	Mouse G-protein co
40	431.5	20.2	361	15 AAR54080	Epstein Barr virus
41	431.5	20.2	361	19 AAW53623	Epstein Barr virus
42	431.5	20.2	361	21 AAY90630	Human G protein-co
43	429	20.1	374	22 AAE04390	Turkey P2Y nucleot
44	428.5	20.1	361	21 AAY90664	Human mutant G pro
45	407	19.1	322	15 AAR48715	G-protein coupled

ALIGNMENTS

RESULT 1
AAW51405
ID AAW51405 standard; Protein; 369 AA.
XX AC
XX AAW51405;
XX DT 12-OCT-1998 (first entry)
XX Mouse protease-activated receptor 3 (PAR3).
XX DE
XX DE
XX KW Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
KW G-protein coupled receptor; agonist; antagonist; thrombosis;
KW atherosclerosis; restenosis; inflammation; blood coagulation;
KW blood clotting; heart attack; stroke; wound healing;
KW adult respiratory distress syndrome; glomerulosclerosis.
XX Mus sp.
XX WO9818456-A1.
XX PD 07-MAY-1998.
XX PF 29-OCT-1997; 97WO-US19732.
XX PR 30-OCT-1996; 96US-0742440.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Connolly A, Coughlin SR, Ishihara H;
XX DR WPI; 1998-271905/24.
XX DR N-PSDB; AAV07372.

PT DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.

PS Claim 3; Page 39-40; 74pp; English.

XX This polypeptide comprises mouse protease-activated receptor 3
 CC (PAR3), a cell surface protein which is specifically activated by
 CC thrombin or a thrombin agonist, thereby activating signalling
 CC events such as phosphoinositide hydrolysis, calcium ion efflux and
 CC platelet aggregation. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAV07372). The human PAR3 amino acid
 CC sequence (see AAW51406) is also provided. Also claimed are vectors,
 CC host cells and an assay device. Host cells are used to screen
 CC compounds for their ability to act as agonists or antagonists of
 CC the effects of thrombin-PAR3 interaction. Agonists are used to
 CC treat wounds, thrombosis, atherosclerosis, restenosis, inflammation
 CC and other thrombin activated disorders. Antagonists (see AAW51415-21)
 CC are used to control blood coagulation and thereby to treat heart
 CC attack and stroke. They also mediate inflammatory and proliferative
 CC responses to injury as occur in wound healing, atherosclerosis,
 CC restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis.

XX Sequence 369 AA;

Query Match 90.8%; Score 1940; DB 19; Length 369;
 Best Local Similarity 100.0%; Pred. No. 2e-199;
 Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 MKLILVAAGLLFLPVTVCQGINVSDNSAKPTLTIKSFNGGPQNTFEEPLSDIEGWTG 76
 Db 1 MKLILVAAGLLFLPVTVCQGINVSDNSAKPTLTIKSFNGGPQNTFEEPLSDIEGWTG 60

Qy 77 ATTITKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSL 136
 Db 61 ATTITKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNSIVTLWKLSL 120

Qy 137 RTKSISLIVFHTNLAIADLLFCVTLPPKIAVHLNNGNWFGEVMCRITTVFVGNMYCAI 196
 Db 121 RTKSISLIVFHTNLAIADLLFCVTLPPKIAVHLNNGNWFGEVMCRITTVFVGNMYCAI 180

Qy 197 LILTCMGINRYLATAPFTYOKLPKRSFSLMCGIVVMVFLYMLPFVILKQEVHLVHSE 256
 Db 181 LILTCMGINRYLATAPFTYOKLPKRSFSLMCGIVVMVFLYMLPFVILKQEVHLVHSE 240

Qy 257 ITTCHDVVDACESPSSRFYFVSLAFEGFLIPVILIFCYTTLIHKLKSKDRILWLYIK 316
 Db 241 ITTCHDVVDACESPSSRFYFVSLAFEGFLIPVILIFCYTTLIHKLKSKDRILWLYIK 300

Qy 317 AVLLILVIFTCFAPTNIILVIHANYHHNTDSLYFMYLIAICLGSLSNCLDPFLYFVM 376
 Db 301 AVLLILVIFTCFAPTNIILVIHANYHHNTDSLYFMYLIAICLGSLSNCLDPFLYFVM 360

Qy 377 SKVVDQLNP 385
 Db 361 SKVVDQLNP 369

RESULT 2

AAW51406

ID AAW51406 standard; Protein; 374 AA.

AC AAW51406;

XX 12-OCT-1998 (first entry)

DE Human protease-activated receptor 3 (PAR3).

XX Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
 KW G-protein coupled receptor; agonist; antagonist; thrombosis;
 KW atherosclerosis; restenosis; inflammation; blood coagulation;
 KW blood clotting; heart attack; stroke; wound healing;
 KW adult respiratory distress syndrome; glomerulosclerosis.

XX Homo sapiens.
 OS Key
 XX Modified-site 25..27 Location/Qualifiers
 FT /note= "Asn is N-glycosylated"
 FT 38..39
 FT Cleavage-site
 FT /note= "thrombin cleavage site"
 FT 48..51
 FT Peptide
 FT /note= "hirudin-like sequence"
 FT 82..84
 FT Modified-site
 FT /note= "Asn is N-glycosylated"
 FT 95..117
 FT Domain
 FT /label= TM1
 FT /note= "transmembrane domain 1"
 FT 126..149
 FT Domain
 FT /label= TM2
 FT /note= "transmembrane domain 2"
 FT 168..191
 FT Domain
 FT /label= TM3
 FT /note= "transmembrane domain 3"
 FT 207..231
 FT Domain
 FT /label= TM4
 FT /note= "transmembrane domain 4"
 FT 261..286
 FT Domain
 FT /label= TM5
 FT /note= "transmembrane domain 5"
 FT 301..323
 FT Domain
 FT /label= TM6
 FT /note= "transmembrane domain 6"
 FT 331..333
 FT Modified-site
 FT /note= "Asn is N-glycosylated"
 FT 335..360
 FT Domain
 FT /label= TM7
 FT /note= "transmembrane domain 7"
 FT 361..385
 XX WO9818456-A1.
 XX 07-MAY-1998.
 XX 29-OCT-1997; 97WO-US19732.
 XX 30-OCT-1996; 96US-0742440.
 XX (REGC) UNIV CALIFORNIA.
 XX Connolly A, Coughlin SR, Ishihara H;
 XX WPI; 1998-271905/24.
 XX N-PSDB; AAV07374.
 XX DNA encoding protease-activated receptor 3 - for detection of
 XX specific agonists and antagonists, potentially useful for treating
 XX e.g. thrombosis, atherosclerosis, inflammation etc.
 XX Claim 3; Page 41-42; 74pp; English.
 XX This polypeptide comprises human protease-activated receptor 3
 CC (PAR3), a cell surface protein which is specifically activated by
 CC thrombin or a thrombin agonist, thereby activating signalling
 CC events such as phosphoinositide hydrolysis, calcium ion efflux and
 CC platelet aggregation. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAV07372), and shows homology to human PAR1
 CC and PAR2 (see AAW51407 and AAW51408). The mouse PAR3 amino acid
 CC sequence (see AAW51405) is also provided. Also claimed are vectors,
 CC host cells and an assay device. Host cells are used to screen
 CC compounds for their ability to act as agonists or antagonists of
 CC the effects of thrombin-PAR3 interaction. Agonists are used to
 CC treat wounds, thrombosis, atherosclerosis, restenosis, inflammation
 CC and other thrombin activated disorders. Antagonists (see AAW51415-21)
 CC are used to control blood coagulation and thereby to treat heart
 CC attack and stroke. They also mediate inflammatory and proliferative
 CC responses to injury as occur in wound healing, atherosclerosis,

CC restenosis, pulmonary inflammation (ARDS) and glomerulosclerosis.

XX Sequence 374 AA;
SQ
Query Match 66.7%; Score 1425.5; DB 19; Length 374;
Best Local Similarity 71.6%; Pred. No. 3.5e-144;
Matches 260; Conservative 44; Mismatches 58; Indels 1; Gaps 1;
QY 17 MKILILVAAGLLFLPVTVCOSGI-NVSDNSAKPTLTIKSFNGGPGQNTFEBFPLSDIEGWT 75
DB 1 MKALIPAAAGLLLLPTFCQSGMENDTNLAKPTLPKTRGAPPNSFEFFPSALEGWT 60
QY 76 GATTIKABCPEDSISLTHVNNATIGYRLSLSSTQVPIAIYILLFVGVGPNVITLWKL 135
DB 61 GATITVKIKPEESASHLVKNATMGYLTSLSTKLPIAIYLLVGVGVPANAVTLMLLF 120
QY 136 LRKTSISLVIFHTNLATADLLFCVTLFPKIAHYHLNGNNWVFGVMCRITTVVFGNMYCA 195
DB 121 FRRSICITTVFYTNLADELFCVTLFPKIAHYHLNGNNWVFGVLCRATTIVIFGNMYCS 180
QY 196 ILILTCMGINRYLATAHPFTYQKLPKRSFSLLMCGIYVVMVFLYMLPFVILKOEYHLVHS 255
DB 181 ILLACISINRYLAIVHPFTYRGIPKHTYALVTCGLWATVFLYMLPFVILKOEYHLVQ 240
QY 256 EITTCDDVDACSPSSFRFYFVSLAPFGFLIPFVIIIFCYTTLIHKLKSKDRIMLYGI 315
DB 241 DIITTCDDVHNTCESSPPQLYFISLAPFGFLIPFVIIIFCYTTLIHKLKSKDRIMLYGI 300
QY 316 KAVLLIIVITICFAPNTIILVIHHANYYYHNTDLSYFMVLIALLCLGSLNSCLDPFLYFV 375
DB 301 KASLLIIVITICFAPNTIILVIHHANYYYHNTDLSYFMVLIALLCLGSLNSCLDPFLYFV 360
QY 376 MSK 378
DB 361 MSK 363

RESULT 3

AA066922
ID AAR66922 standard; Protein; 399 AA.
XX
AC AAR66922;
XX
DT 22-AUG-1995 (first entry)
XX
DE Murine C140 receptor deduced from cDNA.
XX
KW G-protein-coupled receptor; G-protein; C140 receptor.
XX
OS Mus musculus.
XX
PN WO9503318-A.
XX
PD 02-FEB-1995.
XX
PF 26-JUL-1994; 94WO-US08536.
XX
PR 26-JUL-1993; 93US-0057938.
XX
PA (CORT-) COR THERAPEUTICS.
XX
PI Scarbrough RM, Sundelin J;
XX
DR WPI; 1995-075182/10.
DR N-PSDB; AAQ84559.
XX
PT New DNA encoding recombinant C140 receptor - and novel agonists
PT and antagonists and specific antibodies with therapeutic and
PT diagnostic applications.
XX
PS Example; Fig 10; 57pp; English.
XX
CC A cDNA library from a mouse stomach was constructed in lambda gt10

CC and screened with a probe encompassing the C140 genomic clone (see
CC AAQ84557). A single phage clone was isolated and cut with EcoRI. The
CC insert was cloned in pBluescript and pSG5 and sequenced. The
CC complete nt sequence and deduced AA sequence is given in AAQ84559 &
CC AAR66922. 5' RACE resulted in the addition of only 27 bps to the 5'
CC end. The 5' end of the apparent coding region differs from the 5'
CC end of the ORF of genomic DNA; it is believed that the 5' end of
CC the cDNA sequence is correct.

XX Sequence 399 AA;
SQ

Query Match 27.2%; Score 581.5; DB 16; Length 399;
Best Local Similarity 37.6%; Pred. No. 1.6e-53;
Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;
QY 95 VNNATTGYLRSSLSSTQVPIAIYILLFVGVGPNVITLWKLSTRT-KSISLVIFHTNLAI 153
DB 63 IDEFSASITGLTKTTVFLPVIIVFVIGLPSNGMALWIFLFTKKHPAVIYMANLALA 122
QY 154 DLLFCVTLFPKIAHYHLNGNNWVFGVMCRITTVVFGNMYCAIILTCMGINRYLATAHP 213
DB 123 DLLSVITWFPKISYHLGNNWVYGEALCKVLIGFFYGNMYCSILFWTCLSVQRYWVIVNP 182
QY 214 FTYQKLPKRSFSLLMCGIYVVMVFLYMLPFVILKOEYHLVHSIITTCDDVDACSPSSF 273
DB 183 MGHPR-KKANIAGVSLAIWLILFVLTIPLYVMKQTIYIPALNITTCDDVLP-EVLVGD 240
QY 274 RYVYFVSLAFAFGFLIPFVIIIFCYTTLI-----HKLKSKDRILWGIKAVLLILV 323
DB 241 MFNYFUSLAIGVFLFALLUTASAYVLMIKTLRSSAMDEHSEKQRQA----ILRIITVLA 296
QY 324 IFTICFAPNTIILVIHHANYYYHNTDLSYFMVLIALLCLGSLNSCLDPFLYFVMSK 378
DB 297 MYFICFAPSNLLLVHVFILIKTQROSHVYALYVALCLSLNSCIDPFVYVYFVSK 351

RESULT 4

AA01954
ID AAW01954 standard; Protein; 399 AA.
XX
AC AAW01954;
XX
DT 02-APR-1997 (first entry)
XX
DE Murine C140 receptor.
XX
KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..31 /note= "the signal peptide differs from that
FT encoded by a genomic DNA sequence for
FT this receptor (see AAW01952), the signal
FT sequence given here is believed to be
FT the correct sequence"
FT Protein 32..399 /note= "mature protein"
XX
XX WO9623225-A1.
XX
XX 01-AUG-1996.
XX
XX 25-JAN-1996; 96WO-US01179.
XX
XX 25-JAN-1995; 95US-0390301.
XX
XX (CORT-) COR THERAPEUTICS INC.
XX
XX Scarbrough RM, Sundelin J;
XX

DR WPI; 1996-362813/36.
 DR N-PSDB; AAT32038.
 XX Vector for expression C140 cell surface receptor in host cell
 PT useful to identify C140 agonist and antagonists, which are
 PT antihypertensives and elevators of blood pressure, respectively
 XX
 PS Example 4; Fig 10A-B; 60pp; English.
 XX
 CC AAW01954 represents the murine C140 receptor (C140R). DNA encoding C140R
 CC may be engineered so as to allow the recombinant expression of C140R in
 CC a suitable host cell, i.e. by removing the native expression-control
 CC sequences and replacing them with control sequences operable in the
 CC host. Such a recombinant receptor can be expressed on the surface of
 CC oocytes, this provides a good assay system for identifying
 CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked
 CC receptor and a member of the "seven-pass" transmembrane receptor
 CC superfamily (peptide chain of the receptor passes through the cell
 CC membrane seven times, producing seven transmembrane regions within the
 CC receptor molecule). The C140 receptor is involved in controlling blood
 CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
 CC signalling from this receptor, resulting in an increase in blood pressure
 CC and are therefore useful in pharmaceuticals for the treatment of
 CC hypotension (low blood pressure). Conversely agonists (see
 CC AAW01914-W01941) of C140 are useful in pharmaceuticals for the treatment
 CC of hypertension (high blood pressure).
 XX
 SQ Sequence 399 AA;

Query Match 27.2%; Score 581.5; DB 17; Length 399;
 Best Local Similarity 37.6%; Pred. No. 1.6e-53;
 Matches 111; Conservative 65; Mismatches 102; Indels 17; Gaps 5;
 QY 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVVGPSNIVTLWKLSLRT-KSISLVIEHTNLAIA 153
 Db 63 IDEFSASILTGLTTFVPIVYIIIVFGLPSNGMALWIFLFTKXKHPAVIYMANLALA 122
 QY 154 DLLFCVTLFPKIAHYHLNGNNVFGVEMCRITTVVFGYNNMYCAILILTCMGINRYLATAHP 213
 Db 123 DLLSVIWFPLKISYHLHGNVNVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 182
 QY 214 FTYQKLPKRSFSLMCGIVVMVFLYMLPFVILKQEVHLVHSITTCCHDVVDACESSPF 273
 Db 183 MGHPK-KKANIAGVSLAIWLLIFLVTIPIYVNMKQTIYIPALNITTCCHDVLPB-EVLVGD 240
 QY 274 RFYFVSLAFGLPIPFVILFICVTLLI-----HKLSKDRILWGLVKAIVLILV 323
 Db 241 MFNYFLSLAIGVFLFPALLTASAYVLMIKLRSAMDEHSEKQRORA----IRLIITVLA 296
 QY 324 IFTICFAPTNIILVIHHANYYYHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378
 Db 297 MYFICFAPSNLLLVVHVFLKIQKQSHVYALYLVALCLSTLNSCIDPFFVYFVSK 351

RESULT 5
 AAR66920
 ID AAR66920 standard; Protein; 395 AA.
 XX
 AC AAR66920;
 XX
 DT 22-AUG-1995 (first entry)
 XX
 DE Murine C140 receptor.
 XX
 KW G-protein-coupled receptor; G-protein; C140 receptor.
 XX
 OS Mus musculus.
 XX
 FH Key
 FT 1..27 Location/Qualifiers
 FT Protein /label= tentative signal sequence
 FT Region 78..100 /label= transmembrane I
 FT

Region 108..128 /label= transmembrane II
 Region 168..191 /label= transmembrane III
 Region 168..191 /label= transmembrane IV
 Region 204..224 /label= transmembrane V
 Region 266..286 /label= transmembrane VI
 Region 304..325 /label= transmembrane VII
 Modified-site 29 /label= Asn linked glycosylation site
 Cleavage-site 34..35 /label= protease receptor cleavage
 Modified-site 200 /label= Asn linked glycosylation site
 WO9503318-A.
 02-FEB-1995.
 26-JUL-1994; 94WO-US08536.
 26-JUL-1993; 93US-0097938.
 (CORT-) COR THERAPEUTICS.
 Scarborough RM, Sundelin J;
 WPI; 1995-075182/10.
 N-PSDB; AAR64557.
 New DNA encoding recombinant C140 receptor - and novel agonists and antagonists and specific antibodies with therapeutic and diagnostic applications.
 Disclosure; Fig 1; 57pp; English.
 A mouse cosmid genomic library (obtd. from Dr R.A. Wetzel, Washington Univ. School of Medicine, St Louis, Missouri) was screened with two 32P-labeled oligos corresp. to bp 190-249 and 742-801 of the bovine substance K receptor cDNA. In one of the clones isolated (C140) the hybridising region was localised to a 3.7 kb PstI fragment. This fragment was subcloned into pBluescript vector. The hybridising and adjacent regions were sequenced. The nt sequence and the deduced AA sequence are given in AAR64557 & AAR69920 respectively.

Query Match 26.9%; Score 574.5; DB 16; Length 395;
 Best Local Similarity 37.3%; Pred. No. 8.8e-53;
 Matches 110; Conservative 66; Mismatches 102; Indels 17; Gaps 5;
 QY 95 VNNATIGYLRSSLSSTQVIPAIIYLLFVVGPSNIVTLWKLSLRT-KSISLVIEHTNLAIA 153
 Db 59 IDEFSASILTGLTTFVPIVYIIIVFGLPSNGMALWIFLFTKXKHPAVIYMANLALA 118
 QY 154 DLLFCVTLFPKIAHYHLNGNNVFGVEMCRITTVVFGYNNMYCAILILTCMGINRYLATAHP 213
 Db 119 DLLSVIWFPLKISYHLHGNVNVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
 QY 214 FTYQKLPKRSFSLMCGIVVMVFLYMLPFVILKQEVHLVHSITTCCHDVVDACESSPF 273
 Db 179 MGHPK-KKANIAGVSLAIWLLIFLVTIPIYVNMKQTIYIPALNITTCCHDVLPB-EVLVGD 236
 QY 274 RFYFVSLAFGLPIPFVILFICVTLLIHKLS-----KDRILWGLVKAIVLILV 323
 Db 237 MFNYFLSLAIGVFLFPALLTASAYVLMIKLRSAMDEHSEKQRORA----IRLIITVLA 292
 QY 324 IFTICFAPTNIILVIHHANYYYHNTDSLYFMYLIALCLGSLNSCLDPPFLYFVMSK 378

Db 293 MYFICRPSNLLLVHVFYLIKTRQSHVYALYLVALCLSTLNSCIDPFVYFVSK 347

RESULT 6
AAW01952
ID AAW01952 standard; Protein; 395 AA.
XX AAW01952;
AC AAW01952;
XX 01-APR-1997 (first entry)
XX Murine C140 receptor, including putative signal sequence.
DE C140 receptor; G-protein linked; coupled; seven pass; agonist;
XX antagonist; hypertension; hypotension; blood pressure.
KW antagonist; hypertension; hypotension; blood pressure.
XX Mus sp.
XX OS
XX Key Location/Qualifiers
FH 1..27
FT Peptide /note= "putative signal peptide, differs from
FT signal peptide encoded by a cDNA clone of
FT this receptor (see AAW01954), the signal
FT sequence given for the cDNA clone is
FT believed to be the correct sequence"
FT Protein 28..395
FT /note= "mature protein"
FT Modified-site 29
FT /note= "potential Asn-linked glycosylation site"
FT Cleavage-site 34..35
FT /note= "putative protease receptor cleavage site"
FT Region 78..100
FT /note= "transmembrane region I"
FT Region 108..128
FT /note= "transmembrane region II"
FT Region 148..169
FT /note= "transmembrane region III"
FT Region 188..210
FT /note= "transmembrane region IV"
FT Modified-site 220
FT /note= "potential Asn-linked glycosylation site"
FT Region 244..264
FT /note= "transmembrane region V"
FT Region 286..306
FT /note= "transmembrane region VI"
FT Region 324..345
FT /note= "transmembrane region VII"
XX WO9623225-A1.
XX 01-AUG-1996.
XX 25-JAN-1996; 96WO-US01179.
XX 25-JAN-1995; 95US-0390301.
XX (CORT-) COR THERAPEUTICS INC.
XX Scarborough RM, Sundelin J;
XX WPI; 1996-362813/36.
XX N-PSDB; AAT32036.
XX Vector for expression C140 cell surface receptor in host cell
PT useful to identify C140 agonist and antagonists, which are
PT antihypertensives and elevators of blood pressure, respectively
XX Example 1; Fig 1A-B; 60pp; English.
XX AAW01952 represents the murine C140 receptor (C140R), including a
CC putative signal peptide (see features table). DNA encoding C140R may be
CC engineered so as to allow the recombinant expression of C140R in a
CC suitable host cell, i.e. by removing the native expression-control

CC sequences and replacing them with control sequences operable in the
CC host. Such a recombinant receptor can be expressed on the surface of
CC oocytes, this provides a good assay system for identifying
CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked
CC receptor and a member of the "seven-pass" transmembrane receptor
CC superfamily (peptide chain of the receptor passes through the cell
CC membrane seven times, producing seven transmembrane regions within the
CC receptor molecule). The C140 receptor is involved in controlling blood
CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
CC signalling from this receptor, resulting in an increase in blood pressure
CC and are therefore useful in pharmaceuticals for the treatment of
CC hypotension (low blood pressure). Conversely agonists (see
CC AAW01914-W01941) of C140 are useful in pharmaceuticals for the treatment
CC of hypertension (high blood pressure).
XX
SQ Sequence 395 AA;
Query Match 26.8%; Score 572.5; DB 17; Length 395;
Best Local Similarity 37.3%; Pred. No. 1.4e-52;
Matches 110; Conservative 65; Mismatches 103; Indels 17; Gaps 5;
QY 95 VNNATIGYLRSSLSQTQVIPAIVYLLFVGVGPSNIVTLWKLRLT-KSISLVIFHTNLIA 153
DB 59 IDEFSASILTGLTTFVLPVWIIIVFVIGLPSNGMALWIFLFTKKKHPAVIYMANLALA 118
QY 154 DLLFCVTLPFKIAYHLNGNNVFGVWCRTTTFVFGNMYCAILLITCMGINRYLATAP 213
DB 119 DLLSVIWFPLKISYHLGNNVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWVIVNP 178
QY 214 FTYQKLPKRSFSLMCGIUVWVFLVWLPFVILKQYHLVHSEITTCDDVDVDAESPSS 273
DB 179 MGHPR-KKANIAGVSSLAIWLLIFLTIPLYNKQTIYIPALNITTCDDVLP-EVLVGD 236
QY 274 RPYFVSLAFPGFLIPFVIIIFCYTTLIHKLKS-----KDRWLGYIKAVLLILV 323
DB 237 MENYFLSLAIGVFLFPALITASNYVLMIKTLSSANDEHSENKRQRA----ILITVLA 292
QY 324 ITTICFAPTNIILVHHANYYYHNTSDLYFMVLIALLCLGSLNSCLDPFLYFVMSK 378
DB 293 MYFICRPSNLLLVHVFYLIKTRQSHVYALYLVALCLSTLNSCIDPFVYFVSK 347

RESULT 7
AAB35641
ID AAB35641 standard; Protein; 397 AA.
XX AAB35641;
XX 19-FEB-2001 (first entry)
XX Human PAR-2 protein.
XX PAR-2; protease activated receptor-2; ECL-2; inflammatory disease;
KW asthma; chronic obstructive pulmonary; arthritis;
KW inflammatory bowel; psoriasis; eczema; multiple sclerosis.
XX Homo sapiens.
XX WO200063371-A1.
XX 26-OCT-2000.
XX 17-APR-2000; 2000WO-GB01455.
XX 15-APR-1999; 99GB-0008513.
XX (UYSO-) UNIV SOUTHAMPTON.
XX Walls AF, Palmer K, Compton SJ, Cairns JA, Gough AC;
PI WPI; 2000-679599/66.
XX Protease activated receptor 2 variants useful for treating inflammatory

PT diseases such as asthma, arthritis and psoriasis, and as hypertensives,
 PT has reduced sensitivity to trypsin -

PS Claim 2; Page 55; 59pp; English.

XX The present invention relates to a variant protease activated
 CC receptor 2 (PAR-2). The invention is useful for identifying an
 CC individual having a polymorphism in the ECL-2 region of one or
 CC both PAR-2 gene alleles. The invention may be used to develop treatments
 CC for inflammatory diseases such as asthma, chronic obstructive
 CC pulmonary diseases, arthritis, inflammatory bowel diseases, psoriasis
 CC and eczema, multiple sclerosis and to raise blood pressure.

XX Sequence 397 AA;

Query Match 26.6%; Score 569; DB 21; Length 397;
 Best Local Similarity 32.7%; Pred. No. 3.5e-52;
 Matches 120; Conservative 73; Mismatches 136; Indels 38; Gaps 8;

Qy 22 LVAAGLLFLPVTVCQGINVSDNSAKPTLTIKSPNGGPQNTFEFFPLSDIEGWTGATTI 81
 Db 9 LLGAAILLAASLSCGTIOGTRSKGRSLIGKVDG-----TSHV---TGKGVTV 55

Qy 82 KAECPEDSISTLHVNNTATIGYRLSSLTQVIPAIIYLLFVVGVPSPNIVTLWKLRLT-KS 140

Db 56 E-----TVFSVDEFSASVLTGKLTTFPLPIVYTVFVVGLPSPNGMALWVFLRTKKX 107

Qy 141 ISLVIFHTLAIADLLFCVTLPPKIAIHLNGNNWVGEVMCRITTVFYGNMYCAIILT 200

Db 108 HPAAIYMANALADLLSVIWFPLPIAIIHNGNNIYGEALCNVLIGFFYGNMYCSILFMT 167

Qy 201 CMGINRYLATAPHTYOKLPKRSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEITTC 260

Db 168 CLSVQRTWVIVNPMGHSR-KNATAIGISLAIWLLIILVPIYVVKQTIFPALNITTC 226

Qy 261 HDVVDAACESPSF---RFYFVSLAPFGFLIPEVIIIFCYTTLIHKLK-----KDRI 310

Db 227 HDVL-----PEQLLVGMDFNYFLSLAIGVLPFAFLTASAYVLMIRLSRSSAMSENSEK 281

Qy 311 WLGVKAVALLIVFTICFAPTNIIIVHANYHYNTDSLYFMYLIACLGINSCLDP 370

Db 282 RKRAIKLIVTVALMYLICFTPSNLLLVVHFLIKSQSQSHVYALYIIVALCLSTLNSCIDP 341

Qy 371 FLYFVMS 377

Db 342 FVYFVS 348

RESULT 8

AAR66921

XX AAR66921 standard; Protein; 398 AA.

AC AAR66921;

XX 22-AUG-1995 (first entry)

DT Human C140 receptor.

DE G-protein-coupled receptor; G-protein; C140 receptor.

XX Homo sapiens.

Key Location/Qualifiers

FT Protein 1..27

FT /label= signal peptide

FT Modified-site 31

FT /label= Asn linked glycosylation site

FT Modified-site 223

FT /label= Asn linked glycosylation site

FT Cleavage-site 37..38

FT /label= protease receptor cleavage site

FT Region 81..103

FT /label= transmembrane I

FT Region 11..132
 FT /label= transmembrane II
 FT Region 150..174
 FT /label= transmembrane III
 FT Region 191..212
 FT /label= transmembrane IV
 FT Region 245..267
 FT /label= transmembrane V
 FT Region 289..309
 FT /label= transmembrane VI
 FT Region 327..348
 FT /label= transmembrane VII
 XX
 XX WO9503318-A.

XX PD 02-FEB-1995.

XX PF 26-JUL-1994; 94WO-US08536.

XX PR 26-JUL-1993; 93US-0097938.

XX PA (CORT-) COR THERAPEUTICS.

XX PI Scarborough RM, Sundelin J;

XX DR WPI; 1995-075182/10.

XX DR N-PSDB; AAQ84558.

XX New DNA encoding recombinant C140 receptor - and novel agonists
 PT and antagonists and specific antibodies with therapeutic and
 PT diagnostic applications.

XX Disclosure; Fig 2; 57pp; English.

XX The availability of genomic DNA encoding the mouse protease

CC C140 receptor (see Q84557) permitted the retrieval of the corresp.

CC human gene. A human genomic library cloned in the vector EMBL3 was

CC screened using the entire coding region of the murine clone as a

CC probe. The recovered human gene including the DNA sequence and the

CC deduced AA sequence are shown in Q84558 & R66921. Subsequent

CC experiments indicated that the human C140 gene is located in the

CC same region of the long arm of chromosome number 5 (5q12-5q13)

CC as has been reported for the human thrombin receptor gene.

XX Sequence 398 AA;

Query Match 26.4%; Score 563; DB 16; Length 398;

Best Local Similarity 33.5%; Pred. No. 1.5e-51;

Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;

Qy 49 TLTIKSFNGGPQNTFEFFPLSDIEGW---TGATTTIKACPEDSISTLHVNNTATIGVLR 105

Db 21 TLVFLSCTGTRSSKGRSLIGKVDGSHVTKGVTV-----TVFSVDEFSASVLTG 72

Qy 106 SLSTQVIPAIYIILLFVVGVPSPNIVTLWKLRLT-KSISLVIFHTLAIADLLFCVTLPPK 164

Db 73 KLTTFELPIVYTVFVVGLPSPNGMALWVFLRTKKGHPAVIYMANLADLLSVIFPLK 132

Qy 165 IAYHLNGNNWVGEVMCRITTVFYGNMYCAIILTCMGINRYLATAPHTYOKLPKRSF 224

Db 133 IAYHIGNNNIYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSR-KKANI 191

Qy 225 SLLMCGIWMVFLYMLPFVILKQEVHLVHSEITTCDDVDVDAACESPSF-----RFYFVS 280

Db 192 AIGISLAIWLLIILVPIYVVKQTIFPALNITTCDDVL-----PEQLLVGMDFNYFLS 246

Qy 281 LAFFGFLIPEVIIIFCYTTLIHKLK-----KDRIWLGVYKAVLLIILVFTICFAPTN 334

Db 247 LAIGVFLFPAFLTASAYVLMIRLSRSSAMSENSEKRAIKLIVTVALMYLICFTPSNL 306

Qy 335 ILVIHANYHYNTDSLYFMYLIACLGINSCLDPFLYFVMS 377

Db 307 LLVHVYFLIKSQSQSHVYALYIIVALCLSTLNSCIDPFPVYFVS 349

RESULT 9
AAW01953
ID AAW01953 standard; Protein; 398 AA.
XX
AC AAW01953;
XX
DT 01-APR-1997 (first entry)
XX
DE Human C140 receptor, with putative signal sequence.
XX
KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
XX antagonist; hypertension; hypotension; blood pressure.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT /note= "putative signal peptide, differs from
FT this peptide encoded by a cDNA clone of
FT this receptor (see AAW01955), the signal
FT sequence given for the cDNA clone is
FT believed to be the correct sequence"
FT
FT Protein 28..398
FT /note= "mature protein"
FT Modified-site 31
FT /note= "potential Asn-linked glycosylation site"
FT
FT Cleavage-site 37..38
FT /note= "putative protease receptor cleavage site"
FT Region 81..103
FT /note= "transmembrane region I"
FT Region 111..132
FT /note= "transmembrane region II"
FT Region 151..174
FT /note= "transmembrane region III"
FT Region 191..212
FT /note= "transmembrane region IV"
FT Modified-site 223
FT /note= "potential Asn-linked glycosylation site"
FT Region 245..267
FT /note= "transmembrane region V"
FT Region 289..309
FT /note= "transmembrane region VI"
FT Region 327..348
FT /note= "transmembrane region VII"
XX
PN W09623225-A1.
XX
XX
PD 01-AUG-1996.
XX
XX 25-JAN-1996; 96WO-US01179.
XX
XX 25-JAN-1995; 95US-0390301.
XX
XX (CORT-) COR THERAPEUTICS INC.
XX
XX Scarborough RM, Sundelin J;
XX
XX WPI; 1996-362813/36.
XX N-PSDB; AAT32037.
XX
XX Vector for expression C140 cell surface receptor in host cell
XX useful to identify C140 agonist and antagonists, which are
XX antihypertensives and elevators of blood pressure, respectively
XX
XX Example 2; Fig 2A-B; 60pp; English.
XX
XX AAW01953 represents the human C140 receptor (C140R), including a
XX putative signal peptide (see features table). DNA encoding C140R may be
XX engineered so as to allow the recombinant expression of C140R in a
XX suitable host cell, i.e. by removing the native expression-control
XX sequences and replacing them with control sequences operable in the

CC host. Such a recombinant receptor can be expressed on the surface of
CC oocytes, this provides a good assay system for identifying
CC agonists/antagonists of C140R. The C140 receptor is a G-protein linked
CC receptor and a member of the "seven-pass" transmembrane receptor
CC superfamily (peptide chain of the receptor passes through the cell
CC membrane seven times, producing seven transmembrane regions within the
CC receptor molecule). The C140 receptor is involved in controlling blood
CC pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
CC signalling from this receptor, resulting in an increase in blood pressure
CC and are therefore useful in pharmaceuticals for the treatment of
CC hypotension (low blood pressure). Conversely agonists (see
CC AAW01914-W01941) of C140 are useful in pharmaceuticals for the treatment
CC of hypertension (high blood pressure).
XX
XX Sequence 398 AA;
XX
Query Match 26.4%; Score 563; DB 17; Length 398;
Best Local Similarity 33.5%; Pred. No. 1.5e-51;
Matches 115; Conservative 72; Mismatches 128; Indels 28; Gaps 7;
QY 49 TLTIKSFNGGPONTPEEFPLSDIEGW--TGATTTIKARCPEDSISTLHVNNATIGYLR 105
DB 21 TLVFLSCTGTNRSSKGRSLIGKVDGTSHTVGKVTVE-----TVFSVDEFSASVLTG 72
QY 106 SLSTQVIPAIIYLLFVGVPSNIVTLWKLRLT-KSISLVFIHTNLAIAADLLFCVTLPFK 164
DB 73 KLTVFLPIVYITIVFVVGSLPNSGMALWFLFRTKKKHPAVYMANLALADLLSVIWFPLK 132
QY 165 IAYHLNGNNVFGVEMCRITTVFYGNMYCAILLTCMGINRYLATAHPTYQKLPKRSF 224
DB 133 IAYHNGNNVYGEALCNVLIGFFYGNMYCSILFMTCLSVQRYWVWINPMGHSR-KKANI 191
QY 225 SLLMCGIVVMVFLVPLVILKQEVHLVHSEITTCDDVVVDACESPSSF----RFYFVS 280
DB 192 AIGISLAIWLLILLVTIPLYVVKQITFIPALMITTCDDVL-----PEQLLVGDMFNYFLS 246
QY 281 LAFFGFLIPFVLIIFCYTTLIHKLKS-----KDRIMLVYKAVLLILVIFTICFAPTNI 334
DB 247 LAIGVFLPFAFLTASAYVLMIRLSSAMDENSEKKRRAIKLIVTVLAMYLICFTPSNL 306
QY 335 ILVVIHHANYYYHNTDSLYFWYLLIALCLGSLNSCLDPLFYFVMS 377
DB 307 LLVVHYFLIKSQGQSHVYALYIVALCLSLNSCIDPFYFVVS 349
RESULT 10
AAW51408
ID AAW51408 standard; Protein; 394 AA.
XX
XX AAW51408;
XX
DT 12-OCT-1998 (first entry)
XX
DE Human protease-activated receptor 2 (PAR2).
XX
KW Protease-activated receptor 2; PAR2; PAR3; thrombin receptor;
XX human.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Cleavage-site 36..37
XX /note= "thrombin cleavage site"
XX
XX W09818456-A1.
XX
XX 07-MAY-1998.
XX
XX 29-OCT-1997; 97WO-US19732.
XX
XX 30-OCT-1996; 96US-0742440.
XX
XX (REGC) UNIV CALIFORNIA.
XX PA

XX Connolly A, Coughlin SR, Ishihara H;
XX WPI; 1998-271905/24.
XX
XX DNA encoding protease-activated receptor 3 - for detection of
XX specific agonists and antagonists, potentially useful for treating
XX e.g. thrombosis, atherosclerosis, inflammation etc.
XX
XX Example 1; Page 43-44; 74pp; English.
XX
XX This polypeptide comprises human protease-activated receptor 2
XX (PAR2). The physiological activator of PAR2 remains unknown;
XX it is not activated by thrombin. The invention relates to novel
XX mouse and human PAR3 (see AAW51405-06) that show homology to PAR2 and
XX which are specific receptors for thrombin. They can be used to
XX screen for specific agonists and antagonists of thrombin useful
XX e.g. for treating atherosclerosis, thrombosis and inflammation.
XX
XX Sequence 394 AA;

Query Match 25.9%; Score 553.5; DB 19; Length 394;
Best Local Similarity 32.2%; Pred. No. 1.6e-50;
Matches 118; Conservative 73; Mismatches 136; Indels 39; Gaps 8;

QY 22 LVAAGLLFLPVTVCOSGINVSDNSAKPTLTIKSPNGGPQNTFEFFPLSDIEGWTGATTI 81
DB 9 LLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
QY 82 KAECPEDSISTLHVNNATIGYLRSSLSQVPIAIYILLFVVGPSNIVTLWKLRLT-KS 140
DB 56 E-----TVFSVDEFSASVLTGKLTTFVLPVIVTVFVVGPSNGMALWVFLFRKK 107
QY 141 ISLVIFHTNLAIADLLFCVTLFPKIAIYHLNGNNVFGVEMCRITTVVFGNMYCAILILT 200
DB 108 HPAIVYMANLALADLLSVIWFPLKIAIYHIGNNWIYGEALCNVLIGFFYGNMYCSILFMT 167
QY 201 CMGINRYLATAHPFTYQKLPKRSFSLMCGIVMMVFLYMLPFVILKQEHVHSEITTC 260
DB 168 CLSVQRWVIVNPMGHSR-KKANIAGISLAIWLLTLVTVPLVYVVKQTFIPALNITTC 226
QY 261 HDVVD---ACESPSSPFYFVSLAFGLIPFVIIIFCYTTLIHKLK-----KDRIW 311
DB 227 HDVLPQELLVGDP-----FLSLAIGVFLPAPLTASAVVLMIRLSSAMSENSEKR 279
QY 312 LGYIKAVLLILVITFCFAPTNIIIVHANYHHNTDSLYFMYLIALCLGSLNSCLDPF 371
DB 280 KRAIKLIVTLAMYLICFTPSNLLLVVHYFLIKSQGSHVYALYIVALCLSTLNSCIDPF 339
QY 372 LYFVMS 377
DB 340 VYFVS 345

RESULT 11
AAW01955
ID AAW66923 standard; Protein; 397 AA.
XX AAW66923;
XX

XX 22-AUG-1995 (first entry)
XX Human C140 receptor encoded by cDNA.
XX G-protein-coupled receptor; G-protein; C140 receptor.
XX Homo sapiens.

XX W09503318-A.
XX 02-FEB-1995.
XX 26-JUL-1994; 94WO-US08536.

XX 26-JUL-1993; 93US-0097938.
XX (CORT-) COR THERAPEUTICS.
XX
XX Scarborough RM, Sundelin J;
XX WPI; 1995-075182/10.
XX N-PSDB; AAQ84560.
XX
XX New DNA encoding recombinant C140 receptor - and novel agonists
XX and antagonists and specific antibodies with therapeutic and
XX diagnostic applications.
XX Example; Fig 11; 57pp; English.
XX
XX A human intestinal tumour cDNA library was subjected to PCR using
XX primers designed from the genomic clone (see AAQ84558) and the
XX amplified fragment was cloned in pSG5 and sequenced. There are
XX four AA differences between the cDNA encoded sequence and that
XX encoded by the genomic DNA. The genomic DNA sequence and deduced
XX AA sequence are given in AAQ84560 & AAR66923.
XX
XX Sequence 397 AA;

Query Match 25.8%; Score 551; DB 16; Length 397;
Best Local Similarity 32.2%; Pred. No. 3e-50;
Matches 118; Conservative 72; Mismatches 139; Indels 38; Gaps 8;

QY 22 LVAAGLLFLPVTVCOSGINVSDNSAKPTLTIKSFNGGPQNTFEFFPLSDIEGWTGATTI 81
DB 9 LLGAAILLAASLSCSGTIQGTNRSSKGRSLIGKVDG-----TSHV---TGKGVTV 55
QY 82 KAECPEDSISTLHVNNATIGYLRSSLSQVPIAIYILLFVVGPSNIVTLWKLRLT-KS 140
DB 56 E-----TVFSVDEFSASVLAGKLTTFVLPVIVTVFVVGPSNGMALWVFLFRKK 107
QY 141 ISLVIFHTNLAIADLLFCVTLFPKIAIYHLNGNNVFGVEMCRITTVVFGNMYCAILILT 200
DB 108 HPAIVYMANLALADLLSVIWFPLKIAIYHIGNNWIYGEALCNVLIGFFYGNMYCSILFMT 167
QY 201 CMGINRYLATAHPFTYQKLPKRSFSLMCGIVMMVFLYMLPFVILKQEHVHSEITTC 260
DB 168 CLSVQRWVIVNPMGHSR-KKANIAGISLAIWLLTLVTVPLVYVVKQTFIPALNITTC 226
QY 261 HDVVDACESPSF-----RFYFVSLAFGLIPFVIIIFCYTTLIHKLK-----KDR 310
DB 227 HDVL-----PEQLLVGDMFNYFLSLAIGVFLPAPLTASAVVLMIRLSSAMSENSEK 281
QY 311 WLGYIKAVLLILVITFCFAPTNIIIVHANYHHNTDSLYFMYLIALCLGSLNSCLDP 370
DB 282 KRAIKLIVTLAMYLICFTPSNLLLVVHYFLIKSQGSHVYALYIVALCLSTLNSCIDP 341
QY 371 FLYFVMS 377
DB 342 FVYFVS 348

RESULT 12
AAW01955
ID AAW01955 standard; Protein; 397 AA.
XX AAW01955;
XX

XX 02-APR-1997 (first entry)
XX Human C140 receptor.
XX C140 receptor; G-protein linked; coupled; seven pass; agonist;
XX antagonist; hypertension; hypotension; blood pressure.
XX Homo sapiens.

CC physiological responses whereas antagonists are used to downregulate
 CC these activities. The PAR4 protein is further useful for dissecting the
 CC effects of thrombin or other activating proteases in the clotting
 CC pathway from the effects of these proteases at the cellular level.
 CC Agonists are specifically useful in promoting the proliferation
 CC and/or differentiation of platelets, in mediating inflammatory events,
 CC responses to vascular injury, chemotaxis or mitogenesis, and in
 CC producing growth factors. Antagonists are useful as research reagents
 CC for characterizing sites of ligand-receptor interaction.
 CC Note: This sequence is not shown in the specification, but is derived
 CC from the sequence of the PAR4 shown in figure 1, and from the
 CC signal/FLAG epitope peptide shown on page 64.
 XX
 SQ Sequence 408 AA;

Query Match 23.9%; Score 509.5; DB 20; Length 408;
 Best Local Similarity 30.5%; Pred. No. 8.8e-46;
 Matches 117; Conservative 74; Mismatches 133; Indels 39; Gaps 7;

QY 12 SQDIKMLILVAAGLFLPVTVCQSGI-----NVSDNSAKPTLTIKS 54
 DB 7 SQKGRLLLLVVSNNL-----LCQGVVSDYKDDDKLEGQTQTPSVYDESG-----S 54

QY 55 FNGQPQNTFEFPLSDIEGWTGATTITIKACPEPDSISTLHVNNATIGYLRSSLSQVIPA 114
 DB 55 TGGGDDSTPILPAP--RGYPGVQV-----CANDS-DTLELPDSSRALLLGMVPTRLVPA 105

QY 115 IYILLFVVGVPNSIVTLKLSLRTKSISLVIFHTNLAIADLLFCVTLPPFKIAYHLNGNNW 174
 DB 106 LYGLVVLVVGIPANGALVVLATQAPLPSTMLMNLATADLLALALPPRIAYHLRGQRW 165

QY 175 VFGVMCRITTVFVGNMYCAILLTCMGINRYLATAPFTYQKPKRSPSLMCGIVVW 234
 DB 166 PFGEACRLATAALYGHMYGVLLLAVALSLDRYLALVPLRARALGRRLALGLCMAWL 225

QY 235 MVFLYMLPFVLKQEYHLVHSEITTCDDVDACESPSPFRFYFVSIAFEGFLIPFVII 294
 DB 226 MAALALPULTQRTFLASDRVLCHDAL-PLDAQASHWQAPFTCLALLGCFLPLLAML 284

QY 295 FCYTLHLKLSKDRIMGLVIKAVLLILVIFTICFAPTNILVIHANYHYHNTDSLYFM 354
 DB 285 LCYGATLHTLAASGRYCHALRTLAVVLASAVFAFFVPSNNLLLLLHYSDPSPSANGNLYGA 344

QY 355 YLIALCGSLNSCLDPLFYFMS 377
 DB 345 YVPSLALSTLNSCVDPIYYVVS 367

RESULT 15

AAR27240
 ID AAR27240 standard; Protein; 425 AA.

XX AC AAR27240;

XX DT 21-MAY-1998 (first entry)

XX DE Human thrombin receptor (TR).

XX KW Diagnosis; cardiovascular disease; wound healing; restenosis;
 KW thrombosis; unstable angina treatment; myocardial infarction;
 KW thrombotic; thromboembolytic stroke.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

FT peptide 1..26

FT peptide /label= sig_peptide

FT peptide 28..425

FT peptide /label= mat_protein

FT peptide 28..41

FT cleavage_site /note= "activation peptide RPESKATNATLDR"

FT cleavage_site 41..42

FT /note= "thrombin-catalysed cleavage-site"

FT disulfide_bond 175..254
 FT /note= "analogous to rhodopsin and beta-2
 FT adrenergic receptor"
 FT domain 100..129
 FT /note= "transmembrane domain I"
 FT domain 137..160
 FT /note= "transmembrane domain II"
 FT domain 177..197
 FT /note= "transmembrane domain III"
 FT domain 219..239
 FT /note= "transmembrane domain IV"
 FT domain 266..294
 FT /note= "transmembrane domain V"
 FT domain 311..335
 FT /note= "transmembrane domain VI"
 FT domain 348..367
 FT /note= "transmembrane domain VII"
 FT region 34..36
 FT /note= "consensus N-linked glycosylation site"
 FT region 62..64
 FT /note= "consensus N-linked glycosylation site"
 FT region 75..77
 FT /note= "consensus N-linked glycosylation site"
 FT region 250..252
 FT /note= "consensus N-linked glycosylation site"
 FT region 259..261
 FT /note= "consensus N-linked glycosylation site"
 XX WO9214750-A.
 XX PN
 XX 03-SEP-1992.
 XX PD
 XX 19-FEB-1992; 92WO-US01312.
 XX PF
 XX 19-FEB-1991; 91US-0657769.
 XX PR 07-NOV-1991; 91US-0789184.
 XX PR
 XX (CORT-) COR THERAPEUTICS INC.
 XX PA (REGC) UNIV CALIFORNIA.
 XX PA
 XX Coughlin SR, Scarborough RM;
 XX PI
 XX WPI; 1992-316119/38.
 XX DR N-PSDB; AAQ28568.
 XX XX
 PT DNA encoding cell surface receptor for thrombin - useful for
 PT determining thrombin in diagnosing e.g. cardiovascular diseases,
 PT also to treat wound healing, restenosis etc.
 XX
 XX Disclosure; Fig 1; 8ipp; English.
 XX PS
 XX The TR DNA sequence can be used in the prepn. of diagnostics to
 CC determine thrombin levels in samples, and screening tools for
 CC candidate substances which affect thrombin activity in vivo.
 CC Thrombosis may be diagnosed in a mammal by measuring the presence,
 CC absence or amt. of the cleaved activation peptide of the TR.
 XX
 SQ Sequence 425 AA;

Query Match 23.8%; Score 509; DB 13; Length 425;
 Best Local Similarity 31.2%; Pred. No. 1.1e-45;
 Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

QY 20 LLIIVAAAGLLFLPVTVCQSGINVDNSAKP-----TLTIKSP-NGGPQNTFEFPLSDI 71
 DB 6 LLLVAAACF-----SLCGPLISARTRRPEESKATNATLDRSFLLRNPNDKYEFFWEDEE 60

QY 72 EGMWTGA-----TTIKAECPEDSISTLHVNNATIGYLRSSLSQVIPAIIYLLFVVGVPNS 127
 DB 61 KNESGLTEYRLVLSINKSSPLQKLPAPFISEDASGYLTSSWLTLPFVPSVYTGVEVSLPLN 120

QY 128 I--VTLWKLSLRTKSISLVIFHTNLAIADLLFCVTLPPFKIAYHLNGNNWVFGVMCRIT 185
 FT /note= "thrombin-catalysed cleavage-site"

Db 121 IMAIVFILKMKVKK-PAVVVMLHLATADVLSVLPFKISYFSGSDMQFSELRCFVT 179
 QY 186 VVFGNMYCAILLTTCGNGRYLATAHP---FTYQKLPKPSFSLMCGIYVWVFLYMLP 242
 Db 180 AAFYCNMYASILEMTVISIDRFVAVVPMQSLWRTLGASFT---CLAIWALAIAGVWP 236
 QY 243 FVILKQEHVHSEITTCDDVDVADACESPSFRFYFVSIAFFGFLIPEFVIIIFCYTTLIH 302
 Db 237 LVLKEQTIQVPLNITTCDDVDVADACESPSFRFYFVSIAFFGFLIPEFVIIIFCYTTLIH 295
 QY 303 KLKS-----KDRILWGLYIKAVLL---ILVIFTICFAPTNIILVIHANYYYH--NTD 349
 Db 296 CLSSSAVANRSKKR-----ALFLSAAVFCIFICGPTNVLIIAHY-SFLSHTSTTE 347
 QY 350 SLYFMYLIAICLGSNCLDPLFYFVMS 377
 Db 348 AAFAYALLCCVCSISSCIDPLIYYAS 375

RESULT 16

AAW51407

ID AAW51407 standard; Protein; 425 AA.

XX AC

XX AAW51407;

XX DT 12-OCT-1998. (first entry)

XX DE Human protease-activated receptor 1 (PAR1).

XX KW Protease-activated receptor 1; PAR1; PAR3; thrombin receptor;

XX KW human.

XX OS Homo sapiens.

XX FH Key

XX Binding-site 52..55

XX FT /note= "thrombin binding site"

XX FT Cleavage-site 41..42

XX FT /note= "thrombin cleavage site"

XX PN WO9818456-A1.

XX PD 07-MAY-1998.

XX PF 29-OCT-1997; 97WO-US19732.

XX PR 30-OCT-1996; 96US-0742440.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Connolly A, Coughlin SR, Ishihara H;

XX DR WPI; 1998-271905/24.

XX DNA encoding protease-activated receptor 3 - for detection of

XX PT specific agonists and antagonists, potentially useful for treating

XX PT e.g. thrombosis, atherosclerosis, inflammation etc.

XX PS Example 1; Page 42-43; 74pp; English.

XX CC This polypeptide comprises human protease-activated receptor 1

XX CC (PAR1), a receptor that mediates thrombin signalling. The

XX CC invention relates to novel mouse and human PAR3 (see AAW51405-06)

XX CC that show homology to PAR1 and which are specific receptors for

XX CC thrombin. They can be used to screen for specific agonists and

XX CC antagonists of thrombin useful e.g. for treating atherosclerosis,

XX CC thrombosis and inflammation.

XX SQ Sequence 425 AA;

Query Match 23.8%; Score 509; DB 19; Length 425;

Best Local Similarity 31.2%; Pred. NO. 1.1e-45;

Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

QY 20 LILVAAGLLFLPVTVCQSGINVDNSAKP-----TLTIKSF-NGGPONTREERPLSDI 71
 Db 6 LLLVAACF-----SLCGPLSARTARRPESKATNATLDPKSLRNPNDKYEPFWEDEE 60
 QY 72 EGMTGAT---TTIKAECPEDSISTLHVNNATIGYLRSLSTQVIPAIIYILLFVVGVPSN 127
 Db 61 KNEISGLTEYRLVYSINKSSPLOKQLPAFISEDASGYLTSSMLTLFVPSVVTGVFVSLPLN 120
 QY 128 I--VTLWKLUSLTKSISLVFHTNIAIADLLFCVTLPFKIAIYHLNNGNNVFGVEMCRIT 185
 Db 121 IMAIVFILKMKVKK-PAVVVMLHLATADVLSVLPFKISYFSGSDMQFSELRCFVT 179
 QY 186 VVFGNMYCAILLTTCGNGRYLATAHP---FTYQKLPKPSFSLMCGIYVWVFLYMLP 242
 Db 180 AAFYCNMYASILEMTVISIDRFVAVVPMQSLWRTLGASFT---CLAIWALAIAGVWP 236
 QY 243 FVILKQEHVHSEITTCDDVDVADACESPSFRFYFVSIAFFGFLIPEFVIIIFCYTTLIH 302
 Db 237 LVLKEQTIQVPLNITTCDDVDVADACESPSFRFYFVSIAFFGFLIPEFVIIIFCYTTLIH 295
 QY 303 KLKS-----KDRILWGLYIKAVLL---ILVIFTICFAPTNIILVIHANYYYH--NTD 349
 Db 296 CLSSSAVANRSKKR-----ALFLSAAVFCIFICGPTNVLIIAHY-SFLSHTSTTE 347
 QY 350 SLYFMYLIAICLGSNCLDPLFYFVMS 377
 Db 348 AAFAYALLCCVCSISSCIDPLIYYAS 375

RESULT 17

AA49570

ID AAY49570 standard; Protein; 425 AA.

XX AC

XX AAY49570;

XX DT 13-JAN-2000 (first entry)

XX DE Human thrombin receptor protein sequence.

XX KW Human; coding sequence polymorphism; vascular pathology gene;

XX KW polymorphic site; phenotype correlation; forensic; paternity testing;

XX KW medicine; genetic analysis; vascular disease.

XX OS Homo sapiens.

XX PN WO950454-A2.

XX PD 07-OCT-1999.

XX PF 26-MAR-1999; 99WO-US06473.

XX PR 01-APR-1998; 98US-0054272.

XX PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX PI Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;

XX DR WPI; 1999-620066/53.

XX DR N-PSDB; AAZ32191.

XX Determination of polymorphisms in genes, especially those identifying

XX predisposition to vascular disease

XX PS Disclosure; Fig 35; 134pp; English.

XX CC AAZ32159 to AAZ32194 represent reference alleles for specifically

XX CC claimed nucleic acid sequences from the present invention which comprise

XX CC polymorphic sites as given in a table in the specification, selected

XX CC from 92 single nucleotide polymorphisms in which the nucleotide at the

XX CC polymorphic site is different from a nucleotide at the same site in a

XX CC reference allele. The nucleic acids, and primers and probes, are used to

XX CC identify polymorphisms, which may predispose an individual to disease.

CC especially a vascular disease. They can also be used in phenotype
CC correlations, forensics, paternity testing, medicine or genetic
CC analysis. AAY49550 to AAY49573 represent the proteins which correspond
CC to some of the reference alleles.

XX
SQ Sequence 425 AA;

Query Match 23.8%; Score 509; DB 20; Length 425;
Best Local Similarity 31.2%; Pred. No. 1.1e-45;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

Qy 20 LILVAAGLLFLPVTVCQSGINVDNSAKP-----TLTIKSP-NGCPONTFEEFPLSDI 71
Db 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDRSFLRNPNKYEPFWEDEE 60

Qy 72 EGMTGAT----TTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNS 127
Db 61 KNEGSLTEYRLVLSINKSPLOKQLPAFISEDAGSLTSSWTLTFVPSVYTGTVFVVSPLN 120

Qy 128 I--VTLWKLSLRTKSIISLVIFHTNLAIADLLFCVTLFPKIAIYHLNNGNNVFGVEMCRITT 185
Db 121 IMAIWFILKMKVKK-PAVVYMLHLATADVLSVLPFKISYFSGSDWQFGSELCRFVT 179

Qy 186 VVFYGNMYCAILIITCMGINRYLATNP---FTYQKLPKPSFLMCGIIVWVFLWMLP 242
Db 180 AAFYCNMYASILLMTVISIDFLAVVYFMOSLWRTLGRASFT---CLAIWALAIAGVWP 236

Qy 243 FVILKQEVHLVHSEITTHDVVDACESSPSFRFYFVSFLAFFGLIPFVLIIFCYTTLIH 302
Db 237 LVLKEQITQVGLMITTHDVNLTEGYYIYFSAFV-FVFLIISTVCVSIIR 295

Qy 303 KLKS-----KDRIMLGIXIKAVLL---ILVIFTICFAPTNIIILVIHHANYVH--NTD 349
Db 296 CLSSSAVANRSKSR-----ALFLSAAVFCIIICGPNVLLIAHY-SFLSHTTTE 347

Qy 350 SLYFWYLIALLCGLSNCLDPLIFVMS 377
Db 348 AAYFAYLLCVCVSSISSCIDPLIYYAS 375

RESULT 18
AAG80697 ID AAG80697 standard; Protein; 425 AA.
XX AC AAG80697;
XX DT 14-MAR-2002 (first entry)
XX DE Human thrombin protein.

KW G-protein coupled receptor; human; analgesic; antirheumatic; nootropic;
KW antiarthritic; neuroprotective; antiparkinsonian; hypotensive; cardiant;
KW antidepressant; antimigraine; cycostatic; antidiabetic; cardiac disorder;
KW muscular activity; immunosuppressive; vasotropic; gene therapy; cancer;
KW cellular processes modulation; intracellular signalling modulation;
KW proliferative disorder; central nervous system disorder; CNS; pain;
KW metabolic disorder; muscular disorder; rheumatoid arthritis; depression;
KW Alzheimer's disease; Parkinson's disease; autoimmune function disorder;
KW hypertension; neuropsychiatric disorder; learning disorder;
KW memory disorder; bipolar effective neurological disorder;
KW diabetes mellitus; autoimmune disorder; thrombin.

XX OS Homo sapiens.

FX Key Location/Qualifiers
FT Domain 102..132
FT /label= TM1
FT /note= "Transmembrane domain 1"
FT Domain 138..165
FT /label= TM2
FT /note= "Transmembrane domain 2"
FT Domain 176..200
FT /label= TM3

FT Domain /note= "Transmembrane domain 3"
FT 215..241
FT /label= TM4
FT /note= "Transmembrane domain 4"
FT 258..296
FT /label= TM5
FT /note= "Transmembrane domain 5"
FT 313..340
FT /label= TM6
FT /note= "Transmembrane domain 6"
FT 347..379
FT /label= TM7
FT /note= "Transmembrane domain 7"
XX WO200196400-A2.
XX 20-DEC-2001.
XX 18-JUN-2001; 2001WO-US19464.
XX 16-JUN-2000; 2000US-212331P.
XX 16-FEB-2001; 2001US-269758P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Glucksmann MA, Silos-Santiago I;
XX WPI; 2002-098057/13.
XX New isolated human G protein-coupled receptor polypeptide, 52871, for
XX diagnosing and treating proliferative, central nervous system, pain,
XX cardiac, metabolic or muscular disorders
XX Disclosure; Page 106-107; 112pp; English.

CC This invention describes a novel isolated G protein-coupled receptor
CC (GPCR) polypeptide (I), designated 52871. The product of the invention
CC has analgesic, antirheumatic, antiarthritic, nootropic, neuroprotective,
CC antiparkinsonian, hypotensive, antidepressant, antimigraine, cycostatic,
CC antidiabetic, immunosuppressive, vasotropic, cardiant, muscular activity
CC and can be used for gene therapy, cellular processes modulation and
CC intracellular signalling modulation. (I), nucleic acid (II) encoding (I)
CC or an antibody derived from (I) is useful in screening assays, predictive
CC medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical
CC trials and pharmacogenomics), and in methods of treatment (e.g.,
CC therapeutic and prophylactic). (I) or (II) are useful for treating
CC 52871-associated disorders such as proliferative disorders, central
CC nervous system disorders, pain or pain disorders, cardiac disorders,
CC metabolic disorders or muscular disorders. (I) or (II) are useful for
CC treating pain disorders such as rheumatoid arthritis, CNS disorder such
CC as Alzheimer's disease, Parkinson's disease, autoimmune function disorders
CC such as hypertension, neuropsychiatric disorders such as depression,
CC learning or memory disorders, bipolar effective neurological disorders
CC such as migraine, cardiac-related disorders such as stenosis, hormonal
CC disorders such as diabetes mellitus, autoimmune disorders and cancer.
CC This sequence represents the human thrombin protein described in the
CC method of the invention.

XX
SQ Sequence 425 AA;

Query Match 23.8%; Score 509; DB 23; Length 425;
Best Local Similarity 31.2%; Pred. No. 1.1e-45;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;

Qy 20 LILVAAGLLFLPVTVCQSGINVDNSAKP-----TLTIKSP-NGCPONTFEEFPLSDI 71
Db 6 LLLVAACF-----SLCGPLLSARTRARRPESKATNATLDRSFLRNPNKYEPFWEDEE 60

Qy 72 EGMTGAT----TTIKAECPEDSISTLHVNNATIGYLRSSLSSTQVIPAIIYLLFVVGVPNS 127
Db 61 KNEGSLTEYRLVLSINKSPLOKQLPAFISEDAGSLTSSWTLTFVPSVYTGTVFVVSPLN 120

Qy 128 I--VTLWKLSLRTKSIISLVIFHTNLAIADLLFCVTLFPKIAIYHLNNGNNVFGVEMCRITT 185

Db 121 IMAIVVFLKMKVKK-PAVVYMLHLATADVLVSVLPFKISYYPSGSDMQFGSELGRFVT 179
Qy 186 VVFGNMYCAILITTCGINRYLATAP-FTYQKLPKRSFSLMCGIIVWVFLYMLP 242
Db 180 AAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASFT---CLAIWALAIAGVVP 236
Qy 243 FVILKQEHVHSEITTHDWDVADCEPSSPRFYFVSFLAFFGLIPFVIIICFTYTLIH 302
Db 237 LVLKEQTIQVGLNITTHDVLNTELLEGYYAYFSAFSAVF-FVPLIISTVCYVSIIR 295
Qy 303 KLKS-----KDRIMLGVIKAVLL---ILVIFTICFAPTNIILVIHANYHH--NTD 349
Db 296 CLSSSAVANRSKSR-----ALFLSAAVFCIIICFGPTNVLIIAHY-SFLSHTSTTE 347
Qy 350 SLYFMYLIALLCLGSLNSCLDPFLFYVMS 377
Db 348 AAYFAYLLCCVSSISSCIDPLIYYAS 375
RESULT 19
ID AAE17032 standard; Protein; 425 AA.
AC AAE17032;
XX 18-APR-2002 (first entry)
DT Human thrombin protein.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 102..132
FT Domain /label= TW1
FT /note= "Transmembrane domain"
FT 138..165
FT Domain /label= TW2
FT /note= "Transmembrane domain"
FT 176..200
FT Domain /label= TW3
FT /note= "Transmembrane domain"
FT 215..241
FT Domain /label= TW4
FT /note= "Transmembrane domain"
FT 268..296
FT Domain /label= TW5
FT /note= "Transmembrane domain"
FT 313..340
FT Domain /label= TW6
FT /note= "Transmembrane domain"
FT 347..379
FT Domain /label= TW7
FT /note= "Transmembrane domain"
FN WO200202602-A2.
PD 10-JAN-2002.
XX 29-JUN-2001; 2001WO-US20751.
XX 30-JUN-2000; 2000US-0608921.
XX (MILL-) MILLENNIUM PHARM INC.
XX Tsai F;
XX

DR WPI; 2002-140083/18.
XX New isolated G protein-coupled receptor nucleic acid molecule, SLGP,
PT useful for diagnosing and treating cellular proliferation, growth,
PT differentiation, or migration disorders e.g. cancer, arthritis and
XX myocardial ischemia
XX Disclosure; Page 158-159; 182pp; English.
XX The present invention relates to an isolated G protein-coupled receptor
CC (GPCR) nucleic acid molecule, SLGP which encodes a polypeptide. SLGP is
CC useful as target or therapeutic agent for diagnosis and treatment of
CC cellular proliferation, growth, differentiation, or migration disorders
CC (e.g., cancer, arthritis, retinal and optic disk neovascularisation, and
CC tissue ischaemia, such as myocardial ischaemia). SLGP is useful in
CC screening assays, detection assays (e.g., chromosomal mapping, tissue
CC typing, forensic biology), predictive medicine (e.g., diagnostic assays,
CC prognostic assays, monitoring clinical trials and pharmacogenomics), and
CC in methods of treatment (e.g., therapeutic and prophylactic). SLGP is
CC useful as reagents or targets in assays applicable to treatment and
CC diagnosis of SLGP-mediated or related disorders. SLGP is useful in gene
CC therapy, to express SLGP protein, to detect SLGP mRNA or a genetic
CC alteration in a SLGP gene, and to modulate SLGP activity. SLGP is useful
CC to map their respective genes on a chromosome, and thus locate gene
CC regions associated with genetic disease, to identify an individual from a
CC minute biological sample (tissue typing), and to aid in forensic
CC identification of a biological sample. The present sequence is human
CC thrombin protein used in the invention.
XX
XX Sequence 425 AA;
Query Match 23.8%; Score 509; DB 23; Length 425;
Best Local Similarity 31.2%; Pred. NO. 1.1e-45;
Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;
Qy 20 LILVAAAGLLFLPVTVCQSGINVDNSAKP-----TLTIKSF-NGGPNQTEFFPLSDI 71
Db 6 LULVAAACF-----SLCGPLLSARTRARRPESKATNATLDPFRLLRNPNDKYEPFWEDEE 60
Qy 72 ECGWTGAT-----TTIKAECPEDSISTLHVNNATIGYLRSLSTQVIPAIVILLFVGVGPSN 127
Db 61 KNEESGLTEYRLVNSINKSSPLQKLPAPISADASGYLTSSWLTFLFVPSVTGTVFVWSLPLN 120
Qy 128 I--VTLWKLRLTKSISLVIEHTNLAIADLLFCVTLPPKIAHYHLNGNNVFGVCMCRIT 185
Db 121 IMAIVVFLKMKVKK-PAVVYMLHLATADVLVSVLPFKISYYPSGSDMQFGSELGRFVT 179
Qy 186 VVFGNMYCAILITTCGINRYLATAP-FTYQKLPKRSFSLMCGIIVWVFLYMLP 242
Db 180 AAFYCNMYASILLMTVISIDRFLAVVYPMQSLSWRTLGRASFT---CLAIWALAIAGVVP 236
Qy 243 FVILKQEHVHSEITTHDWDVADCEPSSPRFYFVSFLAFFGLIPFVIIICFTYTLIH 302
Db 237 LVLKEQTIQVGLNITTHDVLNTELLEGYYAYFSAFSAVF-FVPLIISTVCYVSIIR 295
Qy 303 KLKS-----KDRIMLGVIKAVLL---ILVIFTICFAPTNIILVIHANYHH--NTD 349
Db 296 CLSSSAVANRSKSR-----ALFLSAAVFCIIICFGPTNVLIIAHY-SFLSHTSTTE 347
Qy 350 SLYFMYLIALLCLGSLNSCLDPFLFYVMS 377
Db 348 AAYFAYLLCCVSSISSCIDPLIYYAS 375
RESULT 20
ID AAY45035
XX AAY45035 standard; Protein; 426 AA.
XX AAY45035;
XX 31-MAY-2000 (first entry)
XX Human thrombin receptor, a member of protease activated receptor family.
XX

XX Human; thrombin receptor; Thr; PAR-1; protease activated receptor;
 KW antisenase molecule; PAR antibody; cytostatic; therapeutic;
 KW metastatic tumour cell; placental implantation; invasive cell.
 XX Homo sapiens.
 OS
 XX
 FH Location/Qualifiers
 FT 37..61
 FT /note= "conserved unique region comprising protease
 FT activated domain and hirudin binding domain."
 FT Therapeutically useful antisenase molecules are designed
 FT based on this region"
 FT 42..51
 FT /label= TRAP
 FT /note= "Thrombin-receptor activating peptide used
 FT for producing anti-Thr antibodies"
 FT Misc-difference 426
 FT /note= "Encoded by TAG"
 FT Region 251..257
 FT /note= "conserved residues in second extracellular
 FT loop used for designing therapeutically useful
 FT antisenase molecules "
 XX
 PN W0200008150-A1.
 XX
 XX 17-FEB-2000.
 XX
 XX 05-FEB-1999; 99WO-IL00079.
 XX
 XX 07-AUG-1998; 98IL-0125698.
 XX
 XX (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
 XX
 XX Bar-Shavit R;
 XX
 DR WPI; 2000-205706/18.
 DR N-PSDB; AA250771.
 XX
 XX Treating metastatic tumor cells useful for treating disorders involving
 PT placenta implantation in a female comprises administration of an
 PT antisenase molecule complementary to an RNA sequence of a protease
 PT activated receptor protein -
 XX
 XX Example 2; Fig 1b; 46pp; English.
 XX
 CC The patent discloses a method to treat metastatic tumour cells using
 CC an antisenase molecule comprising a polynucleotide complementary to an
 CC RNA sequence of a protease activated receptor (PAR) protein, or an
 CC antibody capable of binding to a PAR protein. The antisenase molecules and
 CC antibodies of PAR protein are also used to treat disorders associated
 CC with implantation of placenta. The present sequence is a
 CC human thrombin receptor (THR) (also known as PAR-1), which is a
 CC G-coupled protein belonging to the PAR family. This sequence is used to
 CC produce antisenase molecules or antibodies against PAR protein for
 CC therapeutic use.
 XX
 XX Sequence 426 AA;
 SQ
 Query Match 23.8%; Score 509; DB 21; Length 426;
 Best Local Similarity 31.2%; Pred. No. 1.1e-45;
 Matches 121; Conservative 82; Mismatches 137; Indels 48; Gaps 14;
 QY 20 LILVAAGLLFPVTVQSGINSDNSAKP-----TLTKSF-NGGPQNTFEFFPLSDI 71
 Db 6 LLLVAACE-----SLCGFLSARTRARRPESKATNALDPRFLLRNPNKIPFWEDEE 60
 QY 72 EGMTGAT-----TTIKARCPEDSISTLHVNNATIGYLRSSLSQVPIAYILLFVVGVPNS 127
 Db 61 KNESGLTEYRLVLSINKSPLOKQLPAFISEDASGYLTSMLTFVPSVYGVTVFVWSLPLN 120
 QY 128 I--VTLWKLSTRTKISLVIPHTNLAIADLLFCVTLFPFKIAYHLNGNNWVFGVMCKRITT 185

Db 121 IMAIVVFIILMKVKK-PAVVMHLATADVLFVSVLPFKISYFSGSDNQFSGELCRFVT 179
 QY 186 VVFYGNMYCAIILTCWGINRYLATAHP---FTYQKLPKRSPSLLMCGIVWNVFLYMLP 242
 Db 180 AAFYCNMYASILLMTVISIDRFVAVVPMQSLSWRTLGRASFT--CLAIWALAIAGVVP 236
 QY 243 FVILKOEYHLVHSEITTHDHDVADCESPSRFRYYFVSLAFPGFLIPFVHIIICYYTTLIH 302
 Db 237 LVLKEQTIQVPGNLITTHDVLNETLLEGVYAYFSAFSAVF--FFVPLIISTVCYVSIIR 295
 QY 303 KUKS-----KDRWLGVYKAVLL---ILVIFTICFAPTNILIVIHANYYH--NTD 349
 Db 296 CUSSSAVANRSKSR-----ALFLAAVFCIFICFGPTNVLIIAHY-SFLUSHTSTTE 347
 QY 350 SLYFMVLIALLCLGSLNSCLDPFLYFVMS 377
 Db 348 AAYFALLCCVCVSSISSCIDPLIYYVAS 375
 RESULT 21
 AAY15082
 ID AAY15082 standard; Protein; 385 AA.
 XX
 AC AAY15082;
 XX
 DT 25-JAN-2000 (first entry)
 XX
 DE Human protease-activated receptor 4 (PAR4).
 XX
 KW Human protease-activated receptor 4; PAR4; thrombin receptor;
 KW phosphoinositide hydrolysis; calcium efflux; platelet aggregation;
 KW platelet activation; thrombin agonist; thrombin antagonist; therapeutic;
 KW wound; blood coagulation; heart attack; stroke; inflammatory response;
 KW proliferative response; atherosclerosis; restenosis; glomerulosclerosis;
 KW pulmonary inflammation; ARDS; adult respiratory distress syndrome.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 48..53
 FT /label= PAR4_activating_peptide
 FT Cleavage-site 47..48
 FT /label= Thrombin_cleavage_site
 XX
 PN W09943809-A2.
 XX
 PD 02-SEP-1999.
 XX
 XX 11-FEB-1999; 99WO-US02983.
 XX
 PR 27-FEB-1998; 98US-0032397.
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA
 PI Coughlin SR, Kahn M;
 XX
 XX WPI; 1999-619953/53.
 XX
 PT DNA molecules encoding protease-activated receptor 4, useful in
 XX compound assays for thrombin agonist and antagonist activity -
 PS Claim 7; Fig 5; 69pp; English.
 XX
 CC The present sequence is a human protease-activated receptor 4
 CC (PAR4), a G protein-coupled thrombin receptor expressed on cell surface.
 CC PAR4 is activated by thrombin and mediates signalling events e.g.
 CC phosphoinositide hydrolysis, calcium efflux and platelet aggregation.
 CC The receptor is highly expressed in spleen cells and likely to be
 CC involved in thrombin-mediated activation of platelets and other
 CC haematopoietic cells. It is used for screening novel thrombin agonists
 CC or antagonists. The agonists are used as therapeutics to treat wounds,
 CC promote clotting and as reagents to activate platelets in diagnostic
 CC tests. Antagonists are used to control blood coagulation, treat heart

CC attacks and strokes, and block inflammatory and proliferative responses
CC that occur in normal wound healing and variety of diseases including
CC atherosclerosis, restenosis, pulmonary inflammation (ARDS) and
CC glomerulosclerosis.

XX SQ Sequence 385 AA;

Query Match 23.8%; Score 507.5; DB 20; Length 385;
Best Local Similarity 31.4%; Pred. No. 1.3e-45;
Matches 111; Conservative 76; Mismatches 151; Indels 15; Gaps 6;

Qy 27 LLFLPVTVCQSGINVSNSAKPTTIK--SPNGSQPTTFEEFPLSDIEGWTGATTIKAE 84

Db 5 LLLWPLVL--GFSLSGCTQTPSVYDESGSGGGDDSTPSILPAP--RGYPGQV----- 53

Qy 85 CPEDSISTLHNWNTATIGVLRSSLSQVPIPAIYIILLFVGVGPSNIVTWKLSLRKYSISLV 144

Db 54 CANDS-DTLEPLDSRALLLGWPTRLVPALYGLVVLVGLPANGALWVLATQAPRLPST 112

Qy 145 IFHTNLATADLLFCVTLPFKIAYHLNGNNVFGVMCRITTVFYGNMYCAILLITCMGI 204

Db 113 MLLNWLATADLLALAPPRAYHLRGORWPFGEACRLATAALYGHMYSVLLAAVSL 172

Qy 205 NRYLATARPFTYQKPKRSFSLMCGIWMVFLYMLPFVILKQEVHLVHSEIITCTHDVV 264

Db 173 DRYLALVHLPRARALGRRLALGLCMAAWLMAALALPLTLQRTFLARSDRVLCCHDAL 232

Qy 265 DACSPSPFRYYFVSIAFFGLFPFVIIIFCYTTLHLKSKDRIWGLYKAVLLILVI 324

Db 233 -PLDAQSHWQPAFTCLALLGCFPLLAMLLCYGATLHTLAASGRRYGHALRTAVVLAS 291

Qy 325 FTICFAPTNILVIHANYVHNHDSLVFMYLIAALCLGSLNSCLDPFLFYFMS 377

Db 292 AVAFVPSNLLLLHYDPSPSANGNLYGAVPSLSALSTLNSCVDPIFYIYVS 344

RESULT 22

AA50135

ID AA50135 standard; Protein; 385 AA.

XX AC AA50135;

DT 31-JAN-2000 (first entry)

XX DE Human protease-activated receptor PAR4.

XX KW Protease-activated receptor; PAR4; ZCHEMR2; G protein coupled;
KW cellular signalling; protease; tethered ligand; N-terminal;
KW proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist;
KW antagonist; cellular response; physiological response; clotting pathway;
KW platelet; proliferation; differentiation; mediation; inflammatory process; vascular injury; chemotaxis; mitogenesis;
KW growth factor; production.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Region 1..78 /note= "Extracellular N-terminal region"

FT Peptide 1..17 /note= "Signal peptide"

FT Cleavage-site 17..18 /note= "Cleaved by signal peptidase"

FT Protein 18..385 /note= "Mature non-activated human PAR4"

FT Cleavage-site 47..48 /note= "Cleaved by thrombin or trypsin to activate"

FT Protein 48..385 /note= "Activated human PAR4"

FT Region 48..53 /note= "Tethered hexapeptide activating ligand

FT (AA50140)" Modified-site 56

FT /note= "N-glycosylated"

FT Domain 79..102 /note= "Transmembrane domain 1"
FT Region 103..110 /note= "Intracellular loop 1"
FT Domain 111..132 /note= "Transmembrane domain 2"
FT Region 133..150 /note= "Extracellular loop 1"
FT Domain 151..172 /note= "Transmembrane domain 3"
FT Region 173..191 /note= "Intracellular loop 2"
FT Domain 192..213 /note= "Transmembrane domain 4"
FT Region 214..239 /note= "Extracellular loop 2, determines specificity for activating peptide"
FT Region 228..230 /note= "These three residues are conserved among PAR1-4"
FT Domain 240..263 /note= "Transmembrane domain 5"
FT Region 264..283 /note= "Intracellular loop 3"
FT Domain 284..305 /note= "Transmembrane domain 6"
FT Region 306..316 /note= "Extracellular loop 3"
FT Domain 317..343 /note= "Transmembrane domain 7"
FT Region 344..385 /note= "Intracellular C-terminal region"

WO950415-A2.

XX PN

XX 07-OCT-1999.

XX 31-MAR-1999; 99WO-US07100.

XX 01-APR-1998; 98US-0053866.

XX (ZYMO) ZYMOGENETICS INC.

XX (UNIW) UNIV WASHINGTON.

XX Xu W, Presnell SR, Yee DP, Foster DC;

XX WPI; 1999-633640/54.

XX N-PSDB; AA232747, AA232748.

XX Novel protease activated receptor 4, useful for screening for

XX (antagonists for promoting the proliferation and/or differentiation of

XX platelets and in mediating inflammatory events

XX Claim 8; Fig 1; 85pp; English.

XX This sequence represents a human protease-activated receptor, PAR4 (also

XX referred to as ZCHEMR2). The cDNA was identified from EST (expressed

XX sequence tag) sequences with homology to the three known protease-

XX activated receptors (PAR1, PAR2 and PAR3). Protease-activated receptors

XX (PARs) are a subfamily of G protein coupled receptors which are capable

XX of mediating cellular signalling in response to proteases (e.g.,

XX thrombin). They are characterised by a tethered peptide ligand at the

XX extracellular N-terminus that is generated by proteolysis. PAR4 is

XX activated by thrombin or trypsin cleavage at Arg47/Gly48, which generates

XX a new N-terminus corresponding to the tethered ligand (a hexapeptide).

XX Agonists of PAR4 are useful for upregulating cellular or physiological

XX responses whereas antagonists are used to downregulate these

XX activities. The PAR4 protein is further useful for dissecting the

XX effects of thrombin or other activating proteases in the clotting

XX pathway from the effects of these proteases at the cellular level.

XX Agonists are specifically useful in promoting the proliferation

XX and/or differentiation of platelets, in mediating inflammatory events,

XX responses to vascular injury, chemotaxis or mitogenesis, and in

XX producing growth factors. Antagonists are useful as research reagents

XX

CC for characterising sites of ligand-receptor interaction.

XX SQ Sequence 385 AA;
Query Match 23.8%; Score 507.5; DB 20; Length 385;
Best Local Similarity 31.4%; Pred. No. 1.3e-45;
Matches 111; Conservative 76; Mismatches 151; Indels 15; Gaps 6;
Qy 27 LLFLPVTVCOSGINVSDNSAKPTLTIK--SFNGGPONTFEEFPLSDIEGTGATTIKAE 84
Db 5 LLLWPLVL--GFSLSGGTQTPSVYDESGSTGGDDSTPSILPAP--RGYPGVQV----- 53
Qy 85 CPEDSISTLHVNNATIGYLRSSLSQTQVIPAIIYLLFVVGVPNSIVTLKLSRTKSLV 144
Db 54 CANDS-DTLELPDSSRALLGWPTRLVPAIYGLVGLVGLPANGIALWVLTATQAPRLPST 112
Qy 145 IFHTNLAIADLLFCVTLPEFKIAYHLGNWVFGVEMCRITTVFYGNMYCAILLITCMGI 204
Db 113 MLLMNLATADLLALALPPRIAYHLRGORWPFGEAACRLATAALYGHMYSVLLAAVSL 172
Qy 205 NRYLATAHPPTYOKLPKRSFSLMCGIWMVFLYMLPFVILKOEYHLVHSEITTCDDV 264
Db 173 DRYLALVHPLRARALGRRLALGLCMAAWLMAAALALPLTLQRTQFRLARSDRVLCHDAL 232
Qy 265 DACESPSSFRFYYVSVLAFFGLIPFVIIICYTTLIHKLKSKDRIMGLYKAVLLILVI 324
Db 233 -PLDAQASHWQPAFTCLALLGCFPLLAMLLCYGATLHTLAASGRRYGHALRLTAVVLAS 291
Qy 325 FTICFAPNTIILVIHANYHYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 292 AVAFFVPSNLLLLHYSDPSAWGNLYGAYVPSLALSTLNSCVDPPFIYYVVS 344

RESULT 23

AAAY45036
ID AAY45036 standard; Protein; 385 AA.
XX AC AAY45036;
XX DT 31-MAY-2000 (first entry)
XX DE Human protease activated receptor-4.
XX KW Human; PAR-4; protease activated receptor;
XX KW antisenase molecule; PAR antibody; cytostatic; therapeutic;
XX KW metastatic tumour cell; placental implantation; invasive cell.
XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT Region 43..67
XX FT /note= "conserved unique region comprising protease
XX FT activated domain and hirudin binding domain.
XX FT Therapeutically useful antisenase molecules are designed
XX FT based on this region"
XX FT 228..230
XX FT /note= "conserved residues in second extracellular
XX FT loop used for designing therapeutically useful
XX FT antisenase molecules"

XX FN WO200008150-A1.

XX PD 17-FEB-2000.

XX PF 05-FEB-1999; 99WO-IL00079.

XX PR 07-AUG-1998; 98IL-0125698.

XX PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.

XX PI Bar-Shavit R;

XX DR WPI; 2000-205706/18.

DR N-PSDB; AAZ50775.

XX Treating metastatic tumor cells useful for treating disorders involving
XX placenta implantation in a female comprises administration of an
XX antisenase molecule complementary to an RNA sequence of a protease
XX activated receptor protein -
XX Example 3; Fig 11b; 46pp; English.

XX The patent discloses a method to treat metastatic tumour cells using
XX an antisenase molecule comprising a polynucleotide complementary to an
XX RNA sequence of a protease activated receptor (PAR) protein, or an
XX antibody capable of binding to a PAR protein. The antisenase molecules and
XX antibodies of PAR protein are also used to treat disorders associated
XX with implantation of placenta. The present sequence is a human PAR-4
XX protein used for producing antisenase molecules for treating invasive
XX cells.

XX SQ Sequence 385 AA;

Query Match 23.8%; Score 507.5; DB 21; Length 385;
Best Local Similarity 31.4%; Pred. No. 1.3e-45;
Matches 111; Conservative 76; Mismatches 151; Indels 15; Gaps 6;

Qy 27 LLFLPVTVCOSGINVSDNSAKPTLTIK--SFNGGPONTFEEFPLSDIEGTGATTIKAE 84
Db 5 LLLWPLVL--GFSLSGGTQTPSVYDESGSTGGDDSTPSILPAP--RGYPGVQV----- 53
Qy 85 CPEDSISTLHVNNATIGYLRSSLSQTQVIPAIIYLLFVVGVPNSIVTLKLSRTKSLV 144
Db 54 CANDS-DTLELPDSSRALLGWPTRLVPAIYGLVGLVGLPANGIALWVLTATQAPRLPST 112
Qy 145 IFHTNLAIADLLFCVTLPEFKIAYHLGNWVFGVEMCRITTVFYGNMYCAILLITCMGI 204
Db 113 MLLMNLATADLLALALPPRIAYHLRGORWPFGEAACRLATAALYGHMYSVLLAAVSL 172
Qy 205 NRYLATAHPPTYOKLPKRSFSLMCGIWMVFLYMLPFVILKOEYHLVHSEITTCDDV 264
Db 173 DRYLALVHPLRARALGRRLALGLCMAAWLMAAALALPLTLQRTQFRLARSDRVLCHDAL 232
Qy 265 DACESPSSFRFYYVSVLAFFGLIPFVIIICYTTLIHKLKSKDRIMGLYKAVLLILVI 324
Db 233 -PLDAQASHWQPAFTCLALLGCFPLLAMLLCYGATLHTLAASGRRYGHALRLTAVVLAS 291
Qy 325 FTICFAPNTIILVIHANYHYHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377
Db 292 AVAFFVPSNLLLLHYSDPSAWGNLYGAYVPSLALSTLNSCVDPPFIYYVVS 344

RESULT 24

AAAB47623

ID AAB47623 standard; Protein; 385 AA.

XX AC AAB47623;

XX DT 21-JAN-2002 (first entry)

XX DE Human PAR4.

XX KW Human; protease-activated receptor; PAR4; N-terminal; activated PAR4;
XX KW platelet aggregation; inhibition; tumour cell; proliferation.
XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Peptide 48..53

XX FT /label= N-terminal of activated PAR4

XX PN WO200158930-A1.

XX PD 16-AUG-2001.

XX PF 06-FEB-2001; 2001WO-US03807.

XX 09-FEB-2000; 2000US-0500646.
 PR (ZYMO) ZYMOGENETICS INC.
 PA Baindur N, West RR;
 PI WPI; 2001-656678/75.
 DR N-PSDB; AAH43632.
 XX
 PT Peptides comprising an amino acid sequence are capable of stimulating
 PT protease-activated receptor 4 which are useful in diagnosis and therapy
 PT e.g. inhibiting tumor cell proliferation and stimulating platelet
 PT aggregation -
 XX
 PS Disclosure; Page 58-60; 84pp; English.
 XX
 CC This sequence shows human protease-activated receptor, PAR4. Peptides
 CC derived from, or based on, the N-terminal of activated PAR4 (see
 CC AAB47624-77), are capable of stimulating PAR4. These peptides may be
 CC used to activate PAR4 at lower concentrations than wild type PAR4.
 CC These peptides may be used to stimulate platelet aggregation, and for
 CC inhibiting tumor cell proliferation.
 XX
 SQ Sequence 385 AA;
 Query Match 23.8%; Score 507.5; DB 22; Length 385;
 Best Local Similarity 31.4%; Pred. No. 1.3e-45;
 Matches 111; Conservative 76; Mismatches 151; Indels 15; Gaps 6;
 QY 27 LLELPVTCQSGINVSNAKPTLTIK--SPNGGPQNTFEFFPLSDIEGTGATTTIKAE 84
 DB 5 LLLWPLVL--GFSLSGTQTPSVYDESGTGGGDDSTPSILPAP--RGYPGV----- 53
 QY 85 CPEDSISTLHVNNTATIGVLRSSLSQVPIAIYLLFVGVPSNIVTLWKLRLTKSISLV 144
 DB 54 CANDS-DTLELPDSRALLGWVTRLPALYGLVVLVGLPANGALWLVLTQAPRLPST 112
 QY 145 IFHTNLATADLLFCVTLPEFKIAYHLNGNNWVFGVMCRITTVFYGNMYCAILITCGMI 204
 DB 113 MLLNLTADLLALLALPPIAYHLRGORWFFGGAACRLATAALYGHMYGSLGLAASVL 172
 QY 205 NRYLATAHPFTYQKLPKRSFSLMCGIWMVFLYMLPFVILKQYHLVHSEITTCDDV 264
 DB 173 DRYLALVPLRALRGRRALGLCMAWLMMAALALPLTQRTFRLARSDRVLCDDAL 232
 QY 265 DACESPFRFYFVSLAFFGLPFVILIIICYTTLHLKSKDRIMLGVIKAVLLIIV 324
 DB 233 -PLDAQSHWQPAFTCLALGCLPLFLAMLLCYGATLHTLAASGRRYGHARLRTAVVLAS 291
 QY 325 FTICFAPTNIIIVHANYVHNHNTDSLVFMYLIALCLGSLNSCLDPPFLYFVMS 377
 DB 292 AVAFVPSNLLLLLHYDPSPSANGNLYGAVPSLALSTLNSCVDPPFIYYIVS 344
 RESULT 25
 AAR60698
 ID AAR60698 standard; Protein; 425 AA.
 XX
 AC AAR60698;
 XX
 DT 25-JUN-1995 (first entry)
 XX
 DE Fragment of the human thrombin receptor.
 XX
 KW TR; expression.
 XX
 OS Homo sapiens.
 XX
 PN WO9421789-A.
 XX
 PD 29-SEP-1994.
 XX

PF 28-FEB-1994; 94WO-US02388.
 XX
 PR 25-MAR-1993; 93US-0038662.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Betlach MC, Turner GJ;
 XX
 DR WPI; 1994-317010/39.
 DR N-PSDB; AAQ73590.
 XX
 PT Expression of heterologous proteins in halo-bacteria - using
 PT regulatory and stop sequences from halo-bacteria, pref. the
 PT bacterio-rhodopsin gene.
 XX
 PS Disclosure; Fig 14; 118pp; English.
 XX
 CC The sequence is that of the human thrombin receptor fragment.
 CC This is used to exemplify a new expression vector for producing
 CC heterologous polypeptides in a halobacterial host.
 CC See also AAR60691-9.
 XX
 SQ Sequence 425 AA;
 Query Match 23.7%; Score 507; DB 15; Length 425;
 Best Local Similarity 32.1%; Pred. No. 1.7e-45;
 Matches 115; Conservative 76; Mismatches 131; Indels 36; Gaps 12;
 QY 43 DNSAKPTLTIKSF-NGGPQNTFEFFPLSDIEGTGAT---TTIKAECPEDSISTLHVN 97
 DB 25 DDDVDATLQPSRLLRNPNDEKPFWEDEKESGLTEYRLVLSINKSSPLQKLPAPISE 84
 QY 98 ATTYLRSLSSTQVPIAIYLLFVGVPSNI--VTWKLRLTKSISLVIFHTNLATADL 155
 DB 85 DASGTLTSSWLTTFVPSVYGVFVSVLPLNMAIVFILKMKVKK-PAVYMLHLATADV 143
 QY 156 LFCVTLPEFKIAYHLNGNNWVFGVMCRITTVFYGNMYCAILITCGINRYLATARP-- 213
 DB 144 LFVSVLPFKISYFVSGDMQFSGELCRFVTAAYCNMYASILLMTVISIDRLAVVPMQ 203
 QY 214 -FTYQKLPKRSFSLMCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVDACSPSS 272
 DB 204 SLSWRTLGASFT---CLAIWALAIAGVPLVLEQTIQVPGNLNITTCDDVNETLEGY 260
 QY 273 FRVYFVSLAFFGLPFVILIIICYTTLHLKLS-----KDRIMLGVIKAVLL--I 321
 DB 261 YAYFSAFSAVF-FFVPLIISTVCYVSIIRCLSSAVANRSKSR-----ALFLSAV 312
 QY 322 LVIFTICFAPTNIIIVHANYVHNHNTDSLVFMYLIALCLGSLNSCLDPPFLYFVMS 377
 DB 313 FCIFICFGPTNVLIIAHY-SPLSHTSTTEAAYFAVLLCVCVSSISSCIDPLIYYVAS 369
 RESULT 26
 AAY50138
 ID AAY50138 standard; Protein; 385 AA.
 XX
 AC AAY50138;
 XX
 DT 31-JAN-2000 (first entry)
 XX
 DE Mutant human protease-activated receptor PAR4, R47A.
 XX
 KW Protease-activated receptor; PAR4; G protein coupled;
 KW cellular signalling; protease; tethered ligand; N-terminal;
 KW proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist;
 KW antagonist; cellular response; physiological response; clotting pathway;
 KW platelet; proliferation; differentiation; mediation;
 KW inflammatory process; vascular injury; chemotaxis; mitogenesis;
 KW growth factor; production; mutant.
 XX
 OS Homo sapiens.
 OS Synthetic.

XX FH Key Location/Qualifiers
 FT Region 1..78
 FT Peptide /note= "Extracellular N-terminal region"
 FT Peptide 1..17
 FT Peptide /note= "Signal peptide"
 FT Cleavage-site 17..18
 FT Protein /note= "Cleaved by signal peptidase"
 FT Region 18..385
 FT Region /note= "Mature non-activated human PAR4"
 FT Region 48..53
 FT (AA50140) Modified-site 56
 FT Domain /note= "N-glycosylated"
 FT Region 79..102
 FT Region /note= "Transmembrane domain 1"
 FT Region 103..110
 FT Domain /note= "Intracellular loop 1"
 FT Domain 111..132
 FT Region /note= "Transmembrane domain 2"
 FT Region 133..150
 FT Domain /note= "Extracellular loop 1"
 FT Domain 151..172
 FT Region /note= "Transmembrane domain 3"
 FT Region 173..191
 FT Domain /note= "Intracellular loop 2"
 FT Domain 192..213
 FT Region /note= "Transmembrane domain 4"
 FT Region 214..239
 FT /note= "Extracellular loop 2, determines specificity for activating peptide"
 FT Region 228..230
 FT /note= "These three residues are conserved among PAR1-4"
 FT Domain 240..263
 FT Region /note= "Transmembrane domain 5"
 FT Region 264..283
 FT Domain /note= "Intracellular loop 3"
 FT Region 284..305
 FT Region /note= "Transmembrane domain 6"
 FT Domain 306..316
 FT Region /note= "Extracellular loop 3"
 FT Domain 317..343
 FT Region /note= "Transmembrane domain 7"
 FT Region 344..385
 FT /note= "Intracellular C-terminal region"
 XX WO9950415-A2.
 XX 07-OCT-1999.
 XX 31-MAR-1999; 99WO-US07100.
 XX 01-APR-1998; 98US-0053866.
 XX (ZYMO) ZYMOGENETICS INC.
 XX (UNIW) UNIV WASHINGTON.
 XX Xu W, Presnell SR, Yee DP, Foster DC;
 XX WPI; 1999-633640/54.
 XX Novel protease activated receptor 4, useful for screening for
 XX (ant)agonists for promoting the proliferation and/or differentiation of
 XX platelets and in mediating inflammatory events -
 XX Example 2; Page -; 85pp; English.
 XX This sequence represents a mutant human protease-activated receptor PAR4
 XX R47A, compared with PAR4 mutant R68A (AA50139) in studies to evaluate
 XX the importance of a protease cleavage site between residues 47 and 48.
 XX This mutant PAR4 was unable to be cleaved by thrombin or trypsin, in
 XX contrast to the wild-type (AA50135) or the R68A mutant. This cleavage
 XX activates PAR4, exposing a tethered hexapeptide ligand at the new

CC N-terminus. Protease-activated receptors (PARs) are a subfamily of G
 CC protein coupled receptors which are capable of mediating cellular
 CC signalling in response to proteases. Agonists of PAR4 are useful for
 CC upregulating cellular or physiological responses whereas antagonists
 CC are used to downregulate these activities. The PAR4 protein is
 CC further useful for dissecting the effects of thrombin or other
 CC activating proteases in the clotting pathway from the effects of these
 CC proteases at the cellular level. Agonists are specifically useful in
 CC promoting the proliferation and/or differentiation of platelets, in
 CC mediating inflammatory events, responses to vascular injury, chemotaxis
 CC or mitogenesis, and in producing growth factors. Antagonists are useful
 CC as research reagents for characterising sites of ligand-receptor
 CC interaction.
 CC Note: This sequence is not shown in the specification, but is derived
 CC from the sequence of the PAR4 shown in figure 1.
 XX SQ Sequence 385 AA;
 Query Match 23.7%; Score 506.5; DB 20; Length 385;
 Best Local Similarity 31.4%; Pred. No. 1.7e-45;
 Matches 111; Conservative 76; Mismatches 151; Indels 15; Gaps 6;
 QY 27 LLPLPTVCQSGINVSNSAKPTLTIK--SFNGGPONTPEEPPLSDIEGWTGATTIKAE 84
 DB 5 LLLWPLVL---GFSLSGGTQTSVYDESGSTGGDDSTPSILPAP--AGTPGV----- 53
 QY 85 CPEDSISTLHVNNATIGYLRSSLSITQVIPAIVILLFWGVSPSNIVTLWKLSTKTSISLV 144
 DB 54 CANDS-DTLELPDSSRALLLGWVPTLVLPALYGLVLPANGALWLVATQAPRLPST 112
 QY 145 IFHTNIAIADLFCVTLTPFKIAYHLNGNWNVGEVNCRTTIVVYFNMYCAIILTCMGI 204
 DB 113 MLMLNLTADLLALALPPRIAYHLRGQRWPFGEAAACRLAALYGHMYGSVLLAAVSL 172
 QY 205 NRYLATAHPFTYQKLPKRSFLLMCGIVVMVFLYMLPFVILKQEHVHSEITTCDDVV 264
 DB 173 DRYLALVHPLRARALRGRLALGLCNAAWLMAAALPLTLQRTFLARSDRVLCDDAL 232
 QY 265 DACESPSFRFYFVSLAFFGLIPFVITLIHKLSKDRILWGLYKAVLLILVI 324
 DB 233 -PLDAQASHWQPAFTCTALLGCFLLAMLLCYGATLHTLAASGRYGHRLTAVVLAS 291
 QY 325 FTICFAPTNIILVIHANYYYNHTSDSLYFWYLIALCLGSLNSCLDPFLYFVMS 377
 DB 292 AVAFFVPSNLLLLLHYSDPSAGNLYGAYVPSLALSTLNSCVDPFIYVVS 344
 RESULT 27
 ID AAW16314 standard; Protein; 892 AA.
 XX AAW16314;
 XX 16-AUG-1997 (first entry)
 XX Human thrombin receptor-Yeast G-alpha protein fusion.
 XX G-protein coupled receptor; agonist; antagonist; assay;
 XX G-alpha protein; Gpalp; GPAL gene; thrombin receptor.
 XX Chimaeric Homo sapiens;
 XX Chimaeric Saccharomyces cerevisiae.
 FH Key Location/Qualifiers
 FT Peptide 1..13
 FT /label= Sig_peptide
 FT /note= "alpha-factor (STE2) signal peptide"
 FT Protein 14..892
 FT /label= ThR-GPAL_fusion
 FT Region 14..417
 FT /label= ThR
 FT /note= "thrombin receptor amino acids 22-425"
 FT Peptide 418..420

QY 348 TDSLYFMWYIALCLGSLNSCLDPLFYFVMS 377
Db 327 WGNLYGAVPSLSTLNSCLDPLFYFVMS 356

RESULT 32
AAY84815
ID AAY84815 standard; Protein; 359 AA.
XX AC AAY84815;
XX DT 08-AUG-2000 (first entry)
XX DE A human G-protein coupled receptor designated HG52.
XX KW Human; G-protein coupled receptor; HG52; immune system; thrombin;
KW fibrinogen; fibrin; clotting factor; procoagulant; platelet activation;
KW chemotaxis; mitogenesis.
XX OS Homo sapiens.
XX FN WO200020438-A1.
XX PD 13-APR-2000.
XX PF 29-SEP-1999; 99WO-US22634.
XX PR 02-OCT-1998; 98US-0102959.
XX PA (MERI) MERCK & CO INC.
XX PI Liu Q, McDonald TP, Wang R;
XX DR WPI; 2000-317696/27.
XX DR N-PSDB; AAA14828.

PT New recombinant DNA encoding a G-protein coupled receptor designated
PT HG52 is useful to find modulators of thrombin effects and shows
PT homology to thrombin receptors
XX Claim 6; Fig 2; 39pp; English.

XX The present sequence represents a human G-protein coupled receptor
CC designated HG52. HG52 is a member of the rhodopsin family. The HG52
CC RNA is widely expressed in humans as a transcript of about 4.5 kb,
CC especially in cells of the immune system. The HG52 DNA can be used
CC in chromosomal mapping studies, and to identify individuals carrying
CC a disease-carrying gene. Agonists and antagonists of HG52 will be
CC useful as modulators of the effects of thrombin, including conversion
CC of fibrinogen to fibrin in plasma, activation of clotting factors V,
CC VIII, XIII and protein C, the procoagulant function of platelets and
CC endothelial cells, stimulation of platelet activation, chemotaxis of
CC monocytes and lymphocytes, and mitogenesis of lymphocytes and
CC mesenchymal cells such as vascular smooth muscle cells, fibroblasts
XX and epithelial cells.

SQ Sequence 359 AA;
Query Match 23.1%; Score 493; DB 21; Length 359;
Best Local Similarity 35.7%; Pred. No. 4.3e-44;
Matches 106; Conservative 53; Mismatches 116; Indels 22; Gaps 6;

QY 96 NNATIGYLRSLSTQVPIAYILLFVVGVPNSIVTLWKLSTKXIS-LVIFHTNLAIAD 154
Db 10 DNATLQMLRNPAIAVALPVVYSLVAANSIPGNLSLWVLCRMGPRSPSVIFMNLSTVD 69

QY 155 LLFCVTLFPKXAYHLNGNNWVFGVMCRITTFVYGNMYCAILITCMGIRYLATARPF 214
Db 70 LMLASVLPFQIYYHCRNRHHVFGVLLCNVTVAFVANNYSILMTWTCISVERFLGVLYPL 129

QY 215 TYQKLPKRSFLSLMCGIWWVWVFLVMLPFLVKQYHLVHSEITTCDDVDACSPS--- 271
Db 130 SSKRRRRRYAAACAGTWMLLLLTALSPLARTDITLTPVHALGIITCFDVLKWTMLPSVAM 189

QY 272 --SFREYFVSLAFFGFLIPFVIIIFCYTTLIHKL-----KSDRIWLGYIKAVLLI 321
Db 190 WAVFLFTIFILL---FLIPFVITVACYTATILKLTETEEAHGREQRRRAVGLAAVVLL- 244

QY 322 LVIFTTCFAPTNILVIHHANYHYNTDSLYFMYIALCLGSLNSCLDPLFYFVMSK 378
Db 245 --AFVTCFAPNPFVLLAHIVSRFLFYG-KSYHYHYVYKLTCLSCNLNCLDPEVYFASR 298

RESULT 33
AAG78530
ID AAG78530 standard; Protein; 359 AA.
XX AC AAG78530;
XX DT 12-MAR-2002 (first entry)
XX DE G-protein coupled receptor (designated Paul) amino acid sequence.
XX KW G-protein coupled receptor; paul; 7TM receptor; GPC receptor;
KW human; antibacterial; virucide; fungicide; protozoacide;
KW anti-Human Immunodeficiency Virus; analgesic; cytostatic; antidiabetic;
KW anorectic; antiasthmatic; antiparkinsonian; cardiant; hypertensive;
KW hypotensive; osteopathic; antiangular; antimanic; cerebroprotective;
KW antiulcer; antiallergic; antidepressant; antimigraine; antiemetic;
KW tranquiliser; neuroleptic; neuroprotective; nootropic; anticonvulsant;
KW Human Immunodeficiency virus type 1; HIV-1; HIV-2; pain; cancer;
KW diabetes; obesity; anorexia; bulimia; asthma; Parkinson's disease;
KW acute heart failure; hypotension; hypertension; urinary retention;
KW osteoporosis; angina pectoris; myocardial infarction; stroke; ulcer;
KW allergy; benign prostatic hypertrophy; migraine; vomiting; anxiety;
KW schizophrenia; manic depression; delirium; dementia; mental retardation;
KW dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome.
XX OS Homo sapiens.
XX FN US2001029032-A1.
XX PD 11-OCT-2001.
XX PF 18-DEC-2000; 2000US-0739151.
XX PR 09-OCT-1998; 98US-103789P.
XX PR 06-OCT-1999; 99US-0413534.
XX PA (ZHU)/ ZHU Y.
XX PA (LIX)/ LI X.
XX PA (VAVT)/ VAWTER L.
XX PI Zhu Y, Li X, Vawter L;
XX DR WPI; 2001-647985/74.
XX DR N-PSDB; AA164231.

PT New G-protein coupled receptor polypeptide, referred as Paul and
PT encoding polynucleotide, useful for diagnosing and treating cancers,
PT infections, neurological disorders, diabetes, asthma and identifying
PT modulators
XX Claim 1c; Page 13-14; 15pp; English.

XX The invention relates to an isolated polypeptide, a member of G-protein
CC coupled receptor family of polypeptides, comprising a fully defined
CC sequence of 359 amino acids, its 93% identical sequence, or a polypeptide
CC encoded by a polynucleotide comprising a fully defined sequence of 1080
CC base pairs defined in the specification, or a fragment or variant of it.
CC The activity of the protein of the invention may be described as
CC antibacterial, virucide, fungicide, protozoacide, anti-Human
CC Immunodeficiency Virus, analgesic, cytostatic, antidiabetic, anorectic,
CC antiasthmatic, antiparkinsonian, cardiant, hypertensive, hypotensive,
CC osteopathic, antiangular, antimanic, cerebroprotective, antiulcer,
CC antiallergic, antidepressant, antimigraine, antiemetic, tranquiliser,

CC neuroleptic, neuroprotective, nootropic and anticonvulsant. Polypeptides
 CC and polynucleotides of the invention are useful in diagnosis and in
 CC identifying compounds such as agonists and antagonists which are useful
 CC in therapy. They are also useful for treating diseases, including
 CC infections such as bacterial, fungal, protozoan and viral infections,
 CC particularly infections caused by Human Immunodeficiency virus type 1 or
 CC 2 (HIV-1 or HIV-2), pain, cancers, diabetes, obesity, anorexia, bulimia,
 CC asthma, Parkinson's disease, acute heart failure, hypotension,
 CC hypertension, urinary retention, osteoporosis, angina pectoris,
 CC myocardial infarction, stroke, ulcers, allergies, benign prostatic
 CC hyperplasia, migraine, vomiting, psychotic and neurological disorders,
 CC including anxiety, schizophrenia, manic depression, depression, delirium,
 CC dementia, and severe mental retardation, dyskinesias such as Huntington's
 CC disease or Gilles de la Tourette's syndrome. Polypeptides and
 CC polynucleotides of the invention may also be useful as diagnostic
 CC reagents, for example in the detection of mutations in the associated
 CC gene, chromosome localisation studies and expression pattern
 CC determination. The current sequence represents a G-protein coupled
 CC receptor (designated "Paul" in the specification) amino acid sequence.
 XX
 SQ Sequence 359 AA;

Query Match 23.1%; Score 493; DB 22; Length 359;
 Best Local Similarity 35.7%; Pred. NO. 4.3e-44;
 Matches 106; Conservative 53; Mismatches 116; Indels 22; Gaps 6;
 QY 96 NNATIGVLRSSLSQTQVIPAIIYLLFVVGVPNSIVTLWKLRLTKSIS-LVIFHTNLAIAD 154
 DB 10 DNATLQMLRNPAIATVALPVVYSLVAASIPGNLFSVLWLCRRMGPRSPSIVFMINLSVTD 69
 QY 155 LLFCVTLTPFKIAYHLNGNNWVFGVMCRITTVVFGYNNYCAIILTCMGINRYLATAHPF 214
 DB 70 LMLASVLPFQIYYHCNRHHWVFGVLLCNVTVAYANNYSILTMTCISVERFLGVLYPL 129
 QY 215 TYQKLPKRSPSLMCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVVDACESPS--- 271
 DB 130 SSKWRERRRYAVAAACAGTWLTLTALSPARTDLYTPVHALGIITCFDVLKWTMLPSVAM 189
 QY 272 --SFRFYFVSLAFFGLIPFVLIIFCYTTLIHKL-----KSKDRIMWGIKAVLLI 321
 DB 190 WAVEFTIFILL----FLIPFVITVACVTATILKLRTEBAHQRRRAVGLAAVLL- 244
 QY 322 LVIFTCFAPTNIIIVIHANNYYHNTDSLYFMVLIALLCLGSLNSCLDPPFLYFWSK 378
 DB 245 --AFVTCFAPNFFVLAHIVSRLEFG-KSYHVYVYKTLCLSLNCLDPPFYFASR 298

RESULT 34
 AAG80966
 ID AAG80966 standard; Protein; 359 AA.
 AC AAG80966;
 XX
 XX
 DT 28-AUG-2001 (first entry)
 DE Human nGPCR5 #2.
 XX

KW G protein-coupled receptor; nGPCR; seven transmembrane receptor;
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;
 KW cardiovascular disease; proliferative disorder; hormonal disorder;
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;
 KW attention deficit-hyperactivity disorder; attention deficit disorder;
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;
 KW neuroprotective.

OS Homo sapiens.
 XX
 XX WO200136473-A2.
 PN
 XX
 PD 25-MAY-2001.
 XX
 XX

PF 16-NOV-2000; 2000WO-US31581.
 XX
 PR 16-NOV-1999; 99US-0165838.
 PR 17-NOV-1999; 99US-0166071.
 PR 19-NOV-1999; 99US-0166678.
 PR 28-DEC-1999; 99US-0173396.
 PR 22-FEB-2000; 2000US-0184129.
 PR 28-FEB-2000; 2000US-0185421.
 PR 28-FEB-2000; 2000US-0185554.
 PR 02-MAR-2000; 2000US-0186530.
 PR 03-MAR-2000; 2000US-0186811.
 PR 09-MAR-2000; 2000US-0188114.
 PR 17-MAR-2000; 2000US-0190310.
 PR 21-MAR-2000; 2000US-0190800.
 PR 20-APR-2000; 2000US-0198568.
 PR 02-MAY-2000; 2000US-0201190.
 PR 08-MAY-2000; 2000US-0203111.
 PR 25-MAY-2000; 2000US-0207094.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;
 PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;
 XX WPI: 2001-389826/41.
 DR N-PSDB; AAHS1006.
 XX
 XX New G protein-coupled receptor (nGPCR-x) and its encoding
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -
 PT Claim 37; Page 88; 261pp; English.
 PS
 CC The present invention relates to novel G protein-coupled receptors
 CC (nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present
 CC sequence is one such G protein-coupled receptor. GPCRs are also known as
 CC seven transmembrane receptors and function in signal transduction. The
 CC nGPCRx coding sequences are useful for screening a human to diagnose a
 CC disorder affecting the brain or a genetic predisposition, specifically
 CC schizophrenia. nGPCRx are useful for identifying compounds useful for
 CC treating schizophrenia. Detection of nGPCRx in a sample is useful as a
 CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal
 CC failure, rheumatoid arthritis, CNS disorders, infectious such as HIV-1,
 CC metabolic and cardiovascular diseases, proliferative disorders and
 CC hormonal disorders. Modulators of nGPCRx activity have the utility for
 CC treating neurological disorders, including schizophrenia, ADHD/ADD
 CC (attention deficit-hyperactivity disorder/attention deficit disorder),
 CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,
 CC migraine and senile dementia. Additional disorders include inflammatory
 CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune
 CC disorders, cancers, respiratory ailments such as asthma, and inflammatory
 CC diseases e.g. inflammatory bowel disease.
 XX
 SQ Sequence 359 AA;
 Query Match 23.1%; Score 493; DB 22; Length 359;
 Best Local Similarity 35.7%; Pred. NO. 4.3e-44;
 Matches 106; Conservative 53; Mismatches 116; Indels 22; Gaps 6;
 QY 96 NNATIGVLRSSLSQTQVIPAIIYLLFVVGVPNSIVTLWKLRLTKSIS-LVIFHTNLAIAD 154
 DB 10 DNATLQMLRNPAIATVALPVVYSLVAASIPGNLFSVLWLCRRMGPRSPSIVFMINLSVTD 69
 QY 155 LLFCVTLTPFKIAYHLNGNNWVFGVMCRITTVVFGYNNYCAIILTCMGINRYLATAHPF 214
 DB 70 LMLASVLPFQIYYHCNRHHWVFGVLLCNVTVAYANNYSILTMTCISVERFLGVLYPL 129
 QY 215 TYQKLPKRSPSLMCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVVDACESPS--- 271
 DB 130 SSKWRERRRYAVAAACAGTWLTLTALSPARTDLYTPVHALGIITCFDVLKWTMLPSVAM 189
 QY 272 --SFRFYFVSLAFFGLIPFVLIIFCYTTLIHKL-----KSKDRIMWGIKAVLLI 321
 DB 190 WAVEFTIFILL----FLIPFVITVACVTATILKLRTEBAHQRRRAVGLAAVLL- 244
 QY 322 LVIFTCFAPTNIIIVIHANNYYHNTDSLYFMVLIALLCLGSLNSCLDPPFLYFWSK 378
 DB 245 --AFVTCFAPNFFVLAHIVSRLEFG-KSYHVYVYKTLCLSLNCLDPPFYFASR 298

CC the polypeptides activity in patients having a disorder involving CD8 or
CC CD4 T cells or CD34+ bone marrow cells, especially where the disorder
CC involves thrombocytopenia or inflammation. The modulators can
CC be used to treat disorders involving the spleen (e.g. splenomegaly),
CC lung (e.g. adult respiratory distress syndrome), colon (e.g. bacterial
CC enterocolitis), liver (e.g. hepatic injury), the uterus and
CC endometrium, brain, T-cells, skin, heart, blood vessels, red cells,
CC thymus, B-cells, kidney, breast, testis and epididymis, prostate,
CC thyroid, skeletal muscle, pancreas, small intestine, reduced platelet
CC number, precursor T-cell neoplasms, CD3, CD4, and CD8 T lymphocytes.
XX
SQ Sequence 359 AA;

Query Match 22.9%; Score 489; DB 21; Length 359;
Best Local Similarity 34.7%; Pred. No. 1.2e-43;
Matches 102; Conservative 55; Mismatches 121; Indels 16; Gaps 5;
QY 96 NNATIGYLRSSLSQTQVPAIYILLFVGVFSPSNIVTLWKLSTRTKSI-S-LVIFHTNLAIAD 154
DB 10 DNATQLMRNPAIAVALPVVYSLVAASIPGNLSLWVLCRRMGPRSPSVIFMINLSVTD 69
QY 155 LLFCVTLPFKTAYHLNGNNVFGVEMCRITTVVYGNMYCAILLTCGNGINRYLATAHPF 214
DB 70 LMLASVLPFQIYXHCNRHHWFGVLLCNVTVVAFYANNYSILMTCSVERPLGVLYPL 129
QY 215 TYQKLPKRSFSLMCGIWMVYMLPFLVKQYHLVHSEITTCDDVVDACSPS- 271
DB 130 SSKWRRRRYAVACAGTWLILLTALSPLARTDLYPHALGIITCFDLVLTWMLPSVAM 189
QY 272 --SRFYVYVSLAFPGFLIPFVITFCYTTLIHLKSKDRIW-----LGVIKAVLLILVI 324
DB 190 WAVEFLTIFILL-----FLIPPVITVACVTATILKLRTEEAHGREQSAAGVLA 245
QY 325 FTICFAPNTIIVTHANNYYHNTDSLYFMVLIALLCLGSLNSCLDPFLYFVMSK 378
DB 246 FVTCFAPNPNVLLAHIVSRLEFYG-KSYVHYVYKLTCLCLNCLDPFVYFASR 298

RESULT 37
ABG35298
ID ABG35298 standard; Protein; 402 AA.
XX
AC ABG35298;
DT 15-JUL-2002 (first entry)
DE Human PAR1 type thrombin receptor delta 1-49TR.
XX
KW Human; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW inflammatory disease; cell proliferative disease.
XX
OS Homo sapiens.
XX
FN JP2002010784-A.
XX
PD 15-JAN-2002.
XX
PF 29-JUN-2000; 2000JP-0196514.
XX
PR 29-JUN-2000; 2000JP-0196514.
XX
PA (TEIJ) TEIJIN LTD.
XX
DR WPI; 2002-321520/36.
DR N-PSDB; ABK70887.
XX
PT An inhibitor of cell growth mediated by thrombin used to treat
XX inflammatory and cell proliferative diseases -
PS Disclosure; Page 22-24; 44pp; Japanese.
XX
CC The invention relates to a polypeptide or a compound which can inhibit
CC cell growth caused by thrombin. The polypeptide/compound combines to a

CC specific region of the structure of PAR1 type human thrombin receptor
CC participating to cell growth. Preferably, the compound contains the
CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
CC amino acid or peptide sequence). Also included are a modified PAR1 type
CC thrombin receptor gene or its fragment used for obtaining the above
CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a PAR1 type thrombin receptor (or a modified
XX version).
XX
SQ Sequence 402 AA;

Query Match 22.7%; Score 485.5; DB 23; Length 402;
Best Local Similarity 31.8%; Pred. No. 3.2e-43;
Matches 107; Conservative 74; Mismatches 121; Indels 35; Gaps 11;
QY 63 FEEFPLSDIEGWTGAT-----TTIKACPEPDSISTLHVNNTATIGYLRSSLSQTQVPAIYIL 118
DB 29 YEFWEDEEKNESEGLTEYRLVSVINKSPLOKQLPAFISEDASGYLTSSWTLFVPSVYTG 88
QY 119 LFVGVGPSNI--VTWKLSTRTKSI-SLVIFHTNLAIADLLFCVTLPFKIAYHLNGNNVWF 176
DB 89 FVVVSLPLNIMAIWVILKMKVKK-PAVVVMLHLATADVLSVLPFKISYFSGSDWQF 147
QY 177 GEVWCRTITTVFYGNMYCAILLTCGNGINRYLATAHP-----FTYQKLPKRSFSLMCGIWM 233
DB 148 GSELCHFVTAAFYCNMYASILLMTVISIDRFLLAVVYPMOSLSWRTIGRASFT---CLAIW 204
QY 234 VMVFLMPLPFLVKQYHLVHSEITTCDDVVDACSPSPFRFYVYVSLAFPGFLIPFVII 293
DB 205 ALAIAGVVPLLLKEQITQVPLNITTCDDVNLLEGGYAYVYFSAFSAVF-FVPLIIS 263
QY 294 IFCYTTLIHLKKS-----KDRIWGIKAVLL---ILVITFCFAPNTIILVIHAN 342
DB 264 TVCVYSIIIRCLSSSAVANRSKSR-----ALFLSAAVFCIFICGPTNVLIIAHY-S 315
QY 343 YYVH--NTDSLYFMVLIALLCLGSLNSCLDPFLYFVMS 377
DB 316 FLSHTTTEAAYFAYILLVCVSSISCCIDPLIYYAS 352

RESULT 38
ABG35299
ID ABG35299 standard; Protein; 371 AA.
XX
AC ABG35299;
DT 15-JUL-2002 (first entry)
DE Human PAR1 type thrombin receptor delta 1-80.
XX
KW Human; PAR1; thrombin receptor; antiinflammatory; cytostatic;
KW inflammatory disease; cell proliferative disease.
XX
OS Homo sapiens.
XX
FN JP2002010784-A.
XX
PD 15-JAN-2002.
XX
PF 29-JUN-2000; 2000JP-0196514.
XX
PR 29-JUN-2000; 2000JP-0196514.
XX
PA (TEIJ) TEIJIN LTD.
XX
DR WPI; 2002-321520/36.
DR N-PSDB; ABK70888.
XX
PT An inhibitor of cell growth mediated by thrombin used to treat
XX inflammatory and cell proliferative diseases -

XX Disclosure; Page 24-25; 44pp; Japanese.

XX The invention relates to a polypeptide or a compound which can inhibit

CC cell growth caused by thrombin. The polypeptide/compound combines to a

CC specific region of the structure of PAR1 type human thrombin receptor

CC participating to cell growth. Preferably, the compound contains the

CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type

CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional

CC amino acid or peptide sequence). Also included are a modified PAR1 type

CC thrombin receptor gene or its fragment used for obtaining the above

CC polypeptide, a human PAR1 type thrombin receptor protein and its

CC encoding DNA comprising a fully. The polypeptide or the compound is used

CC to treat inflammatory diseases and cell proliferative diseases. The

CC present sequence is a PAR1 type thrombin receptor (or a modified

CC version).

XX Sequence 371 AA;

XX

Query Match 22.7%; Score 484.5; DB 23; Length 371;

Best Local Similarity 34.2%; Pred. No. 3.7e-43;

Matches 101; Conservative 67; Mismatches 96; Indels 31; Gaps 10;

QY 101 GYLSSLSSTQVPAIYILLFVGVPSNI--VTLKLSLRKTSISLVIFHTNLATADLLFC 158

DB 40 GYLTSWLTUFPVSVYTGTVFVSVPLIMAIWVFLKMKVKK-PAVVMYMLLATADVLV 98

QY 159 VTLFPKAYHLNGNNWVFGVMCRITTVFVGNMYCAILITCMGINRYLATAPHP---FT 215

DB 99 SVLPFKISYFSGSDWQFSGELCFVTAFAFCNNYASILLMTVISIDRFVAVYPMQSL 158

QY 216 YQKLPKRSFSLLMCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVDACSPSFRF 275

DB 159 WRTLGASFT---CLAIWALAIAAGVPLLLKEQTIQVPLNITTCDDVNLLEGGYAY 215

QY 276 YFVSLAFFGFLIPFVIIICFTYTLIHKLK-----KDRIWGLYIKAVLL---TLVI 324

DB 216 YFSAFSAVF-FFVPLIISTVCYSIIICLSSSAVANRSKSR-----ALFLSAAVFCI 267

QY 325 FTICFAPTNIILVHHANYHH--NTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377

DB 268 FIICFGFTNVLIIAHY-SFLSHTSTTEAAYFAYLLCVSVSSICCDPLIYYAS 321

RESULT 39

AAW69598

ID AAW69598 standard; Protein; 359 AA.

XX

AC AAW69598;

XX

XX 16-OCT-1998 (first entry)

DT

DE Mouse G-protein coupled receptor 69A08 #2 protein.

XX

XX Mouse; CKDLR201.1; chemokine; G-protein coupled receptor; GPCR;

KW 7 transmembrane receptor; inflammation; asthma; antiviral;

KW abnormal cell proliferation; regeneration; degeneration; atrophy.

XX

OS Mus sp.

XX

XX WO9831810-A2.

XX

XX 23-JUL-1998.

PD

XX

XX 20-JAN-1998; 98WO-US00218.

PF

XX

XX 21-JAN-1997; 97US-0786624.

PR

XX (SCHE) SCHERING CORP.

PA

XX Capone M, Gorman DM, Hedrick JA, Huffine CF, Rossi DL;

PI Vicari A, Zlotnik A;

XX

DR WPI; 1998-414108/35.

XX N-PSDB; AAV40373.

PT Pure or recombinant chemokine CKDLR201.1 - useful, e.g. for treating

XX inflammation and as antiviral agents

PS Claim 2; Page 68-69; 77pp; English.

XX

CC The present sequence is mouse G-protein coupled receptor 69A08 #2

CC protein which comprises a plurality of epitopes. Host cells containing

CC vectors comprising a nucleotide sequence encoding the protein are used

CC to produce recombinant protein. Treatment of a cell, particularly neuron,

CC macrophage or lymphocyte, carrying a G-protein coupled receptor (GPCR),

CC or a receptor responsive to CKDLR201.1 chemokine protein, with an

CC (ant)agonist is used to control physiological development, e.g.

CC alteration of calcium ion influx, a chemoattractant response, morphology,

CC phosphoinositide lipid turnover or an antiviral response. Nucleotide

CC sequences the CKDLR201.1 protein, or GPCRs, are useful as primers or

CC probes, e.g. for detecting and isolating related sequences and for

CC expressing antigenic peptides. Antibodies (Ab) directed against the

CC CKDLR201.1 protein and GPCRs are used to detect or purify the proteins;

CC diagnostically (e.g. for developmental abnormalities); in screening for

CC potential drugs; to inhibit chemokine/receptor activation; (when coupled

CC to a toxin or radioisotope) for killing specific cells, and to raise

CC anti-idiotype antibodies. CKDLR201.1 protein and GPCRs and compounds

CC which bind them can be used to treat inflammation, e.g. asthma; as

CC antiviral agents, and to treat abnormal cell proliferation, regeneration,

CC degeneration and atrophy. Therapeutic agents are administered orally, by

CC injection and rectally.

XX

SQ Sequence 359 AA;

XX

Query Match 20.8%; Score 444.5; DB 19; Length 359;

Best Local Similarity 33.3%; Pred. No. 7e-39;

Matches 93; Conservative 57; Mismatches 110; Indels 19; Gaps 4;

QY 108 STQVIPAIIYILLFVGVPSNIIVTLKLSLRKTSISLVIFHTNLATADLLFCVTLPKFIAY 167

DB 51 SPQLPALYGLVAVGLPANGLAWLATRVPLRPLSTILLMLNLAVADLLALVLPRLAY 110

QY 168 HLGNNWVFGVMCRITTVFVGNMYCAILITCMGINRYLATAPHPYQKLPKRSFSL 227

DB 111 HLRGQWPFGEEAACRVATAALYGHMVGSLVLLAAVSLDRYLVLPURALRGQRUTTG 170

QY 228 MCGIWMVFLYMLPFVILKQYHLVHSEITTCDDVDAC-----ESPSSFRFYF 278

DB 171 LCLVAVLSAATLALPLTLHRQNFLL-APIACC--VMMRCPWLSRPTGERSAWLSWA 227

QY 279 VSLAFFGFLIPFVIIICFTYTLIHKLKSKDRIWGLYIKAVLLILVITTCFAPTNIILVI 338

DB 228 ASL-----PLLAMGLCYGTTTTERALAANGQRYSHALRLTALVLSAVASTPFSNVLLVL 280

QY 339 HHANYHHNTDSLYFMYLIALCLGSLNSCLDPFLYFVMS 377

DB 281 HYSNPSPEAMGNLYGAYVPSLALSTLNSCVDPIFYIYVS 319

RESULT 40

AAW54080

ID AAW54080 standard; Protein; 361 AA.

XX

AC AAR54080;

XX

XX 03-FEB-1995 (first entry)

DT

DE Epstein Barr virus induced (EBI-2) polypeptide.

XX

XX Epstein Barr virus; EBV; induction; detection; diagnosis;

KW lymphocytes; antigen; growth; differentiation; mediator;

KW infectious mononucleosis.

XX

XX Homo sapiens.

OS

